

STIC-Biotech/ChemLib

135240

From: Slobodyansky, Elizabeth
Sent: Friday, October 15, 2004 10:26 PM
To: STIC-Biotech/ChemLib
Subject: 09/712,768

Please search for case 09/712,768:

88

SEQ ID NOs: 1 and 2 against commercial ~~and interference~~ databases.

Thank you.

Elizabeth Slobodyansky, PhD

Primary Examiner

Art Unit 1652
REM 2D83
571-272-0941
MAILBOX 2C70

RECEIVED
OCT 18 2004
TECH/INTEL DIVISION
(STIC)

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2- _____
Date Searcher Picked up: 10/18/04
Date Completed: 10/23/04
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search 1
NA Sequence: # _____
AA Sequence: # 1
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: Q1/Q28
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)

GenCore version S.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 22, 2004, 12:24:15 ; Search time 7311 Seconds

(without alignments)
10827.925 Million cell updates/sec

Title: US-09-712-768a-1

Perfect score: 1674

Sequence: 1 atggagagcgccgcatc.....acaagcatccctgcgactaa 1674

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: gb_ba:*

2: gb_ntg:*

3: gb_in:*

4: gb_cm:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1674	100.0	1674	6 AX113986	Sequence
2	1674	100.0	1674	6 BD015007	Cytochrome
3	1091.4	85.2	2629	1 PDCYCA	Y07533 P.denitrifi
4	1064.8	60.0	2177	1 RSCOXS1	X62645 Rhodobacter
5	991	59.2	3211	1 PDCOX2	X05829 P.denitrifi
6	754.4	45.1	286550	1 SMES91785	AL591785 Sinorhizo
7	750.8	44.9	10790	1 AE009044	AE009044 Agrobacte
8	750.2	44.8	11188	1 AE006001	AE006001 Caulobact
9	740.4	44.2	10835	1 AE008010	AE008010 Agrobacte
10	738.2	44.1	10154	1 AE014356	AE014356 Brucella
11	736.6	44.0	11404	1 AE009583	AE009583 Brucella
12	719.6	43.0	346510	1 AP003011	AP003011 Mesorhizo
13	709.2	42.4	4380	1 NWCOXABC	X89566 N.winograd
14	708.8	42.3	1830	1 BJCXAG	X54318 B.japonicum
15	708.8	42.3	2000	1 BJCXCA	X54800 B.japonicum
16	708.8	42.3	8121	1 BJA242592	AJ242592 Bradyrhiz
17	708.8	42.3	31495	1 BJU33883	U33883 Bradyrhizob
18	708.8	42.3	300900	1 AP005939	AP005939 Bradyrhiz
19	685.8	41.0	349260	1 BX572595	BX572595 Rhodopsu

20 654 39.1 1707 1 RLC0XA
21 577.2 34.5 196050 1 AL646058
22 549.8 32.8 252260 1 BX640436
23 549.8 32.8 343473 1 BX640451
24 545 32.6 257701 1 BX640422
25 520 31.1 12391 1 AE012038
26 515.6 30.8 12829 1 AE004449
27 510.6 30.5 300774 1 AE016912
28 508.8 30.4 12849 1 AE012504
29 492.4 29.4 303226 1 AE016774
30 461 27.5 1341 8 PRU389607
31 458.2 27.4 1341 8 PRU389606
32 457.8 27.3 1341 8 PAU389608
33 456.6 27.3 1406 8 AF020565
34 454.4 27.1 1412 8 AF020564
35 452.8 27.0 12207 1 AE008617
36 450.6 26.9 4815 8 AF009436
37 450.6 26.9 221853 8 MIRSCOX1
38 450.6 26.9 221853 8 AP006444
39 449 26.8 1341 8 PC0389609
40 449 26.8 1413 8 AF193955
41 449 26.8 2875 8 AY300014
42 449 26.8 6244 8 AY300015
43 449 26.8 66924 8 MIRGENA_3
44 447.6 26.7 1415 8 AF020585
45 447.4 26.7 1584 8 PTU77623

ALIGNMENTS

RESULT 1
AX113986
LOCUS AX113986 1674 bp DNA linear PAT 01-MAY-2001
DEFINITION Sequence 1 from Patent EP103603.
ACCESSION AX113986
VERSION AX113986.1 GI:13940146
KEYWORDS Gluconobacter oxydans
SOURCE Gluconobacter oxydans
ORGANISM Gluconobacter oxydans
Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
Acetobacteraceae; Gluconobacter.

REFERENCE 1
AUTHORS Asakura,A., Hoshino,T. and Shinjoh,M.
TITLE Cytochrome c oxidase complex from gluconobacter oxydans
JOURNAL Patent: EP 1103603-A 1 30-MAY-2001,
F. HOFFMANN-LA ROCHE AG (CH)
FEATURES
source
1. 1674
/organism="Gluconobacter oxydans"
/mol_type="unassigned DNA"
/db_xref="taxon:442"
1. 1674
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC37869.1"
/db_xref="GI:13940147"
/translation="MADAIAHGHDEHHEKQFFTRFMSTNHRKIDIGLILYVAGVVGFI
SVLTFVMBLEMDPGVQVYMCLEGLIADASQICTANGHLWNVMVTVGLMMEFVG
IPALFGGNYLMPLOIGAPDVAFPRMNNLSFWLFIAGTAMGVASLPAFGGDLGSG
WGVLYVP;STREAGYSMDLAFVHLSCASSIMGAINMTTFLAMRAFGMDLHKVPL
FNSIITLAILLILALPLVLAGALTMLLDRNEGTTFFNPAGGDDILYQHILWFSGHP
EVIIILPFGIISRVSTFSKRPVGLPFWYAMVAGLGVFWAHMHTVYVMSLT
QQSYFLATWIAVPTGIKFSWIATMGGSVFEPKSPMLWAFGFMFLFTVGGVTSIL
AQGLORADYDYYVVAHYFVNSIGALPAIFAGIYFYMPKESGFAPFEWAALHFWT
FFICANVTFFPHFLGRQMPRIYIDYPEAFALWKNVSYSGAFAPFASFLFFIVIFY
TLVAGRRTRPNPWFGEFADLTLEWTLPSPPAHTFTETLPKRSWDKHPSH"

ORIGIN

Query Match 100.0%; Score 1674; DB 6; Length 1674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGGCAGACGCGCCATTCACGGCCATACACCATGAGAGCAAGGCTTCTTCACGGC 60
 Db
 1 ATGGCAGACGCGCCATTCACGGCCATACACCATGAGAGCAAGGCTTCTTCACGGC 60
 QY 61 TGGTTTCATGTCGACCAACCAACAAAGACATCGGTCTGTATACCTTTAGCGGCTGGT 120
 Db 61 TGGTTTCATGTCGACCAACCAACAAAGACATCGGTCTGTATACCTTTAGCGGCTGGT 120
 QY 121 GTTGGTTTCATGTCGACCAACCAACAAAGACATCGGTCTGTATACCTTTAGCGGCTGGT 180
 Db 121 GTTGGTTTCATGTCGACCAACCAACAAAGACATCGGTCTGTATACCTTTAGCGGCTGGT 180
 QY 181 GTTCAGTACATGTCGACCAACCAACAAAGACATCGGTCTGTATACCTTTAGCGGCTGGT 240
 Db 181 GTTCAGTACATGTCGACCAACCAACAAAGACATCGGTCTGTATACCTTTAGCGGCTGGT 240
 QY 241 CGGACGACACCTGTGGAACGTCATGGTTACCTTACCAATGATATTCGATGATGTTT 300
 Db 241 CGGACGACACCTGTGGAACGTCATGGTTACCTTACCAATGATATTCGATGATGTTT 300
 QY 301 GTGGGTATCCCGCATGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 360
 Db 301 GTGGGTATCCCGCATGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 360
 QY 361 GCTCCGATATGGCTTCCCGCATGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 420
 Db 361 GCTCCGATATGGCTTCCCGCATGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 420
 QY 421 ACCCGATATGGCTTCCCGCATGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 480
 Db 421 ACCCGATATGGCTTCCCGCATGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 480
 QY 481 GTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 540
 Db 481 GTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 540
 QY 541 GCGATTTCCGGTTCATGTTCCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600
 Db 541 GCGATTTCCGGTTCATGTTCCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600
 QY 601 ACACCTTTTGAACATGCG 660
 Db 601 ACACCTTTTGAACATGCG 660
 QY 661 TGGTCGATCTTATACGCGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 720
 Db 661 TGGTCGATCTTATACGCGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 720
 QY 721 ATCACCATGCTGTCGACCGCGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 780
 Db 721 ATCACCATGCTGTCGACCGCGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 780
 QY 781 GGTGACCGGATCTGTACCAACATCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 840
 Db 781 GGTGACCGGATCTGTACCAACATCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 840
 QY 841 ATCATTCTGCGCGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 900
 Db 841 ATCATTCTGCGCGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 900
 QY 901 GTCTTCCGTTACCTGCGGATGCTATGCAATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 960
 Db 901 GTCTTCCGTTACCTGCGGATGCTATGCAATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 960
 QY 961 GTCTGGGCGACCATGTCACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1020
 Db 961 GTCTGGGCGACCATGTCACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1020
 QY 1021 CTGGCCACCATGGTATCGCGGTGGCGACCGGCATTAAGATCTTCTCGTGGATGCCACG 1080
 Db 1021 CTGGCCACCATGGTATCGCGGTGGCGACCGGCATTAAGATCTTCTCGTGGATGCCACG 1080

QY 1081 ATGTGGGGCGGCTCGGTTGAGTTCAAATCGCGATGCTCTGGGCTTTGGCTTTATGTTTC 1140
 Db 1081 ATGTGGGGCGGCTCGGTTGAGTTCAAATCGCGATGCTCTGGGCTTTGGCTTTATGTTTC 1140
 QY 1141 CTGTTTCACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1200
 Db 1141 CTGTTTCACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1200
 QY 1201 TATACGACACCATTTACGTTGGTGGCAGCTTTCATGATGATGATGATGATGATGATGATGATG 1260
 Db 1201 TATACGACACCATTTACGTTGGTGGCAGCTTTCATGATGATGATGATGATGATGATGATGATG 1260
 QY 1261 TTTGGATCTTTCGCGGATCTTACCTTTTACATGCGGAAGTTCTCGGCGCGGCTTTCCCG 1320
 Db 1261 TTTGGATCTTTCGCGGATCTTACCTTTTACATGCGGAAGTTCTCGGCGCGGCTTTCCCG 1320
 QY 1321 GAATGGGCTGCAAGCTGCATCTTCGACCTTCTTCATCGGTGGAAAGTCTTCGCTTCTTC 1380
 Db 1321 GAATGGGCTGCAAGCTGCATCTTCGACCTTCTTCATCGGTGGAAAGTCTTCGCTTCTTC 1380
 QY 1381 CGGACGACCTTCTTCGCGGATCTCAGGATGATGCGCGGCTTACATCGATATCCCGAAGCC 1440
 Db 1381 CGGACGACCTTCTTCGCGGATCTCAGGATGATGCGCGGCTTACATCGATATCCCGAAGCC 1440
 QY 1441 TTCCGCGCTGTGGAACAAAGTCTCGTCTATGATGATGATGATGATGATGATGATGATGATG 1500
 Db 1441 TTCCGCGCTGTGGAACAAAGTCTCGTCTATGATGATGATGATGATGATGATGATGATGATG 1500
 QY 1501 TTTCTATCGTGGTCTTTTGTCTATACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
 Db 1501 TTTCTATCGTGGTCTTTTGTCTATACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
 QY 1561 CCGTGGGCGGAATTCGCGGATCTCGTGGATGATGATGATGATGATGATGATGATGATGATG 1620
 Db 1561 CCGTGGGCGGAATTCGCGGATCTCGTGGATGATGATGATGATGATGATGATGATGATGATG 1620
 QY 1621 ACGTTCGAAACGCTGCCAAGCGCTCGGACTGGGACAAAGCATCCCTCGCACTAA 1674
 Db 1621 ACGTTCGAAACGCTGCCAAGCGCTCGGACTGGGACAAAGCATCCCTCGCACTAA 1674

RESULT 2
 BD015007
 LOCUS
 DEFINITION
 Cytochrome c oxidase complex.
 BD015007
 ACCESSION
 BD015007.1 GI:22555814
 VERSION
 JP 2001169792-A/1
 KEYWORDS
 Gluconobacter oxydans
 ORGANISM
 Gluconobacter oxydans
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
 Acetobacteraceae; Gluconobacter.
 1 (bases 1 to 1674)
 Asakura, A., Hoshino, T. and Araki, M.
 Cytochrome c oxidase complex
 Patent: JP 2001169792-A 1 26-JUN-2001;
 F. HOFMANN LA ROCHE AG
 OS
 Gluconobacter oxydans
 JP 2001169792-A/1
 PD 26-JUN-2001
 PF 17-NOV-2000 JP 2000351502
 PR 17-NOV-1999 EP 9922842.0
 PI AKIRA ASAKURA, TATSUO HOSHINO, MASAKO ARAKI
 PC C12N15/09, C12N1/21, C12N9/08, C12P7/60, C12N15/09, C12R1/01, PC
 C12N1/21, C12R1/01, C12N9/08, C12R1/01, C12P7/60, C12R1/01, PC
 C12N15/00,
 CC (C12N15/00, C12R1/01)
 PC Cytochrome c oxidase complex
 FH Key Location/Qualifiers
 FT CDS Location/Qualifiers
 1..1674
 /organism="Gluconobacter oxydans"

/mol_type="genomic DNA"
/db_xref="taxon:442"

ORIGIN

Query Match 100.0%; Score 1674; DB 6; Length 1674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGACACCGCGCATTCACGCGCATAGACACCATGAGAGCAAGCTTCTTACGGCG 60
DB 1 ATGCGACACCGCGCATTCACGCGCATAGACACCATGAGAGCAAGCTTCTTACGGCG 60
QY 61 TGGTTTCATGTCGACCAACCAACCAAGACATCGTCTCTATACCTTGTAGCGCTGTGT 120
DB 61 TGGTTTCATGTCGACCAACCAACCAAGACATCGTCTCTATACCTTGTAGCGCTGTGT 120
QY 121 GTTGGTTTCATTCCTGCTCTGTTTCAACGCTTACATCGCGCTTGTAGCTGATGATCGGGT 180
DB 121 GTTGGTTTCATTCCTGCTCTGTTTCAACGCTTACATCGCGCTTGTAGCTGATGATCGGGT 180
QY 181 GTTCAGTACATGTCCTTGAAGCGCACGTCCTGATCGCGGATGCTTCGCGACATGTACG 240
DB 181 GTTCAGTACATGTCCTTGAAGCGCACGTCCTGATCGCGGATGCTTCGCGACATGTACG 240
QY 241 GCGAACCGACACCTGTGGAACGTCATGGTTACCTACCATGGTATTCGATGATGATGTTCTTT 300
DB 241 GCGAACCGACACCTGTGGAACGTCATGGTTACCTACCATGGTATTCGATGATGATGTTCTTT 300
QY 301 GTGGTATCCCGCATTTGCTGGTGTGTTTGGTAACTATCTGATCGCGCTGCAATCGCG 360
DB 301 GTGGTATCCCGCATTTGCTGGTGTGTTTGGTAACTATCTGATCGCGCTGCAATCGCG 360
QY 361 GCTCGGATATGCGCTTCCCGGTATGAAACAACTGCTGCTTGTGGCTGTTCATTCGCGGT 420
DB 361 GCTCGGATATGCGCTTCCCGGTATGAAACAACTGCTGCTTGTGGCTGTTCATTCGCGGT 420
QY 421 ACCGCGATGCGCTTGGCTTGGTGTGCTGCAACCGCGGTGAGCGTCAGCTGGTTCGGCG 480
DB 421 ACCGCGATGCGCTTGGCTTGGTGTGCTGCAACCGCGGTGAGCGTCAGCTGGTTCGGCG 480
QY 481 GTTGGTTGGTTCGTACCGCGCTGTGCAACCGCGGTGAGCGTCAGCTGGTTCGTGACCTC 540
DB 481 GTTGGTTGGTTCGTACCGCGCTGTGCAACCGCGGTGAGCGTCAGCTGGTTCGTGACCTC 540
QY 541 GCGATTTTCGGGTTCATTTGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 541 GCGATTTTCGGGTTCATTTGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 601 ACGACCTTCGTGACATGCGCGCCCGCGCATGACGCTGCAACAAAGTGCCTGCTGCTGCG 660
DB 601 ACGACCTTCGTGACATGCGCGCCCGCGCATGACGCTGCAACAAAGTGCCTGCTGCTGCG 660
QY 661 TGGTGCATCTTATACGCGCTTGGCTGATCCTGCTGCGCTGCGCTGCTGCTGCTGCTGCTG 720
DB 661 TGGTGCATCTTATACGCGCTTGGCTGATCCTGCTGCGCTGCGCTGCTGCTGCTGCTGCTG 720
QY 721 ATCACCATGCTGTCACCGACCTAACTTGGCAGACCTTCTTCAATCCTCTGCGCGCG 780
DB 721 ATCACCATGCTGTCACCGACCTAACTTGGCAGACCTTCTTCAATCCTCTGCGCGCG 780
QY 781 GGTGACCGCATTCGTACCAACATCTGCTGCTGCTTGGCAGACCTGCTGCTGCTGCTGCTG 840
DB 781 GGTGACCGCATTCGTACCAACATCTGCTGCTGCTTGGCAGACCTGCTGCTGCTGCTGCTG 840
QY 841 ATCATTTGCTGCGCTTGGCATCATCAGCATGCTGCTGCTGCTTCTGCAAAAGCGCG 900
DB 841 ATCATTTGCTGCGCTTGGCATCATCAGCATGCTGCTGCTGCTTCTGCAAAAGCGCG 900
QY 901 GTCTTCGTTTACCTGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
DB 901 GTCTTCGTTTACCTGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
QY 961 GTCTGGCGCACCATGTATACCGCTTGGTATGCTGCTGACCCGCAATCTTACTTCTATG 1020

RESULT 3
PDCYCA
LOCUS
DEFINITION
PDCYCA 2629 bp DNA linear BCT 23-JAN-1992
cytochrome c oxidase beta subunit I.
Y07533.1 GI:45477
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

961 GTCTGGCGCACCATGTATACACCGTTGGTATGTGCTGACCCAGCAATCTTACTTCTATG 1020
1021 CTGCGCACCATGTATGCTGCGTCCGACCGCATTAAGCATTTCTGCTGGATCGCCACG 1080
1021 CTGCGCACCATGTATGCTGCGTCCGACCGCATTAAGCATTTCTGCTGGATCGCCACG 1080
1081 ATGTGGGGGGCTCGGTTGAGTTCAATCGCGATGCTCTGGGCTTTTGGCTTTATGTTT 1140
1081 ATGTGGGGGGCTCGGTTGAGTTCAATCGCGATGCTCTGGGCTTTTGGCTTTATGTTT 1140
1141 CTGTTCACTGCTGGTGTGACCGGTATGCTGCTGGGCTTGGGCTTGGACCGTGA 1200
1141 CTGTTCACTGCTGGTGTGACCGGTATGCTGCTGGGCTTGGGCTTGGACCGTGA 1200
1201 TATCAGACACCATTTACGTGGGCGATCTTCAATATGATGATGCTGGGCTGCGATC 1260
1201 TATCAGACACCATTTACGTGGGCGATCTTCAATATGATGATGCTGGGCTGCGATC 1260
1261 TTTGCGATCTTCCCGGTATCTACTTTTACATGCGGAGTTCTCGGGCGCGCTTCCCG 1320
1261 TTTGCGATCTTCCCGGTATCTACTTTTACATGCGGAGTTCTCGGGCGCGCTTCCCG 1320
1321 GAATGGGCTGCAAGCTGACCTTCTGGACCTTCTCATCGGTGGGAGAGTCACTGTTCT 1380
1321 GAATGGGCTGCAAGCTGACCTTCTGGACCTTCTCATCGGTGGGAGAGTCACTGTTCT 1380
1381 CCGCAGCAGCTTCTGGGAGCTGAGGATGCTGCGCGCGTTACATCGACTATCCGAGCC 1440
1381 CCGCAGCAGCTTCTGGGAGCTGAGGATGCTGCGCGCGTTACATCGACTATCCGAGCC 1440
1441 TTGCGCTGTGGAAACAAAGTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
1441 TTGCGCTGTGGAAACAAAGTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
1501 TTCTTCATGCTGATCTTCTGCTATACGCTGTTGCTGCGCGCGGAGACCGCTCGAAC 1560
1501 TTCTTCATGCTGATCTTCTGCTATACGCTGTTGCTGCGCGCGGAGACCGCTCGAAC 1560
1561 CCGTGGGGGAAATCCCGCATGCTGGAATGGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
1561 CCGTGGGGGAAATCCCGCATGCTGGAATGGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
1621 ACCTTCGAAACGCTGCCCAAGCGCTCGGACTGGGACAGCATCCCTCGCACTAA 1674
1621 ACCTTCGAAACGCTGCCCAAGCGCTCGGACTGGGACAGCATCCCTCGCACTAA 1674

1 (bases 1 to 2629)
Raitio, M., Pispas, J. M., Metso, T. and Saraste, M.
Are there isoenzymes of cytochrome c oxidase in Paracoccus
denitrificans?
FEBS Lett. 261 (2), 431-435 (1990)
90184495
2155830
2 (bases 1 to 2629)
Raitio, M.
Direct Submission
Submitted (17-APR-1990) M. Raitio, DEPT OF MEDICAL CHEMISTRY, 17,
UNIVERSITY OF HELSINKI, SILTAVUORENPENGER 10, 00170 HELSINKI,
FINLAND

FEATURES	source	Location/Qualifiers	Db
		1. .2629	
		/organism="Paracoccus denitrificans"	
		/mol_type="genomic DNA"	
		/strain="pD122"	
		/sub_strain="ssp. MR-3"	
		/db_xref="taxon:266"	
gene		/clone_lib="EMBL3 lambda"	
		233. .2544	
		/gene="cyc A"	
CDS		233. .700	
		/gene="cyc A"	
		/EC_number="1.9.3.1"	
		/notes="cytochrome-c oxidase"	
		/codon_start=1	
		/transl_table=11	
		/product="cytochrome c 550"	
		/protein_id="CAA68820.1"	
		/db_xref="GI:45478"	
		/db_xref="GOA:P00096"	
		/db_xref="Swiss-Prot:P00096"	
		/translation="MKISYATLAAITLIPAAQDGAAGKEFKFNKCKACHMIQAP DPTDLIKGKTKPMLGYGVGKSIASBEFGYKIGLEVAEKNPDLTWTADLIEYVTD KPEMLVMDTKDGAKTKMTFTXGKNQADVAFLAQNSPDAGGDGEAAAGESN"	
sig_peptide		233. .292	
		/gene="cyc A"	
mat_peptide		293. .2544	
		/gene="cyc A"	
		/note="cytochrome c 550"	
		/EC_number="1.9.3.1"	
		731. .765	
terminator		/gene="cyc A"	
gene		971. .2547	
		/gene="cta DII"	
CDS		971. .2547	
		/product="cytochrome c oxidase beta subunit I"	
		/protein_id="CAA68821.1"	
		/db_xref="GI:45479"	
		/db_xref="GOA:P98002"	
		/db_xref="Swiss-Prot:P98002"	
		/translation="MADAAVHGHDHDTGRFPFRFMFTNKHDIILYLFTAGIVGL ISVGFVYRMELQHPGVQVWCLGELARLADASAGCTPNMGLWNYMIYVHGLVMWFFV YIPALFGFGNYWELPIGAPDMAFPLRNLSYVMYVCGVALGVALIPGNDONGS GVMVLPEPLSTEAGYSMDLAIFAHVSGASSILGALINIIITFLNRPAGNLFKVP LEANSVITAWLILLSLPVLALGAILMLDMRNFQTFPPAGGDPVLYQHILMFPFG PEVTIILPGFGLIISVLTAFKPIFGVLFPMVLAWAIGILGVVVAHHMYTAGMSL TQQAFLMALNTAIVPTGIVFWSIATMMGGSEIEFKTFLMFAFGFLFTFGVGTGVV LSQAPLDRVYDHYVVAHFVWISLVAGVFGIFAGVYVWIGMSGRQYPEWAGQLHFV MFTGSLNLIFFPHQFVIGROCMPPRYIDYVPEFAYWNNISIGAYISFASFLFFIGIVF YTI.FAGKRVNVPVNWNEHADLTLEWTLPSPPPHETPELTPKREDWDRAH"	
terminator		2585. .2615	
ORIGIN			
Query Match	65.2%; Score 1091.4; DB 1; Length 2629;		
Best Local Similarity	78.6%; Pred. No. 1.7e-216;		
Matches 1318; Conservative	0; Mismatches 356; Indels 3; Gaps 1;		
QY	1	ATGGCAGACGGCGCATTCACGGCCA---TGACCACCATGAGACGCAAGGCTCTTTCACG 57	
Ddb	871	ATGGCAGACGGCGGCTTCAGGCCACGGTGACCATCATGACACCGCGGGTCTTCACC 930	
QY	58	CGCTGGTTCATGTGACCAACCAAGACATCGGTCTCTATACCTTGTAGCGGCTGGT 117	
Ddb	931	CGCTGGTTCATGTCAACAAACCAAGGATATCGGTATCCTTTACCTGTTCACGGCGCGC 990	
QY	118	GTTCGTGCTTTCATTTCGGTCCTGTTACCGCTCATCGCTTACGCTGAGTGAATCG 177	

```
QY 1258 ATCTTGGGATCTTGGCGGATCTACTTTACATGCGGAAGTTCTCGGCGCGGCTTTC 1317
D 1259 ATCTTGGGATCTTGGCGGATCTACTTTACATGCGGAAGTTCTCGGCGCGGCTTTC 1317
Db 2131 GTGTTGGGATCTTGGCGGATCTACTTTACATGCGGAAGTTCTCGGCGCGGCTTTC 2190
QY 1318 CCGGAATGGGATCTTGGCGGATCTACTTTACATGCGGAAGTTCTCGGCGCGGCTTTC 1377
D 1319 CCGGAATGGGATCTTGGCGGATCTACTTTACATGCGGAAGTTCTCGGCGCGGCTTTC 1377
Db 2191 CCGGAATGGGATCTTGGCGGATCTACTTTACATGCGGAAGTTCTCGGCGCGGCTTTC 2250
QY 1378 TTCGGGATCTTGGCGGATCTACTTTACATGCGGAAGTTCTCGGCGCGGCTTTC 1437
D 1379 TTCGGGATCTTGGCGGATCTACTTTACATGCGGAAGTTCTCGGCGCGGCTTTC 1437
Db 2251 TTCGGGATCTTGGCGGATCTACTTTACATGCGGAAGTTCTCGGCGCGGCTTTC 2310
QY 1438 GCGTTGGGATCTTGGCGGATCTACTTTACATGCGGAAGTTCTCGGCGCGGCTTTC 1497
D 1439 GCGTTGGGATCTTGGCGGATCTACTTTACATGCGGAAGTTCTCGGCGCGGCTTTC 1497
Db 2311 GCGTTGGGATCTTGGCGGATCTACTTTACATGCGGAAGTTCTCGGCGCGGCTTTC 2370
QY 1498 CTGTTCTTCTGATCTTGGCGGATCTACTTTACATGCGGAAGTTCTCGGCGCGGCTTTC 1557
D 1499 CTGTTCTTCTGATCTTGGCGGATCTACTTTACATGCGGAAGTTCTCGGCGCGGCTTTC 1557
Db 2371 CTGTTCTTCTGATCTTGGCGGATCTACTTTACATGCGGAAGTTCTCGGCGCGGCTTTC 2430
QY 1558 AACCGGTTGGGATCTTGGCGGATCTACTTTACATGCGGAAGTTCTCGGCGCGGCTTTC 1617
D 1559 AACCGGTTGGGATCTTGGCGGATCTACTTTACATGCGGAAGTTCTCGGCGCGGCTTTC 1617
Db 2431 AACCGGTTGGGATCTTGGCGGATCTACTTTACATGCGGAAGTTCTCGGCGCGGCTTTC 2490
QY 1618 CACAGTTCTGAAACCGTCTGCGGATCTACTTTACATGCGGAAGTTCTCGGCGCGGCTTTC 1674
D 1619 CACAGTTCTGAAACCGTCTGCGGATCTACTTTACATGCGGAAGTTCTCGGCGCGGCTTTC 1674
Db 2491 CACAGTTCTGAAACCGTCTGCGGATCTACTTTACATGCGGAAGTTCTCGGCGCGGCTTTC 2547

RESULT 4
LOCUS RSCOX1 2177 bp DNA linear BCT 24-JUL-2002
DEFINITION Rhodobacter sphaeroides gene for cytochrome oxidase subunit 1.
ACCESSION X62645
VERSION X62645.2 GI:21955870
KEYWORDS cytochrome oxidase subunit 1.
SOURCE Rhodobacter sphaeroides
ORGANISM Rhodobacter sphaeroides
Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
Rhodobacteraceae; Rhodobacter.
1 (bases 1 to 2166)
Shapleigh, J.P. and Gennis, R.B.
Cloning, sequencing and deletion from the chromosome of the gene
encoding subunit I of the aa3-type cytochrome c oxidase of
Rhodobacter sphaeroides
Mol. Microbiol. 6 (5), 635-642 (1992)
92204019
1313140
REFERENCE 2 (bases 1 to 2177)
Shapleigh, J.P.
Direct Submission
Submitted (17-OCT-1991) J.P. Shapleigh, University of Illinois,
Dept of Biochemistry, 505 S Mathews, Urbana IL 61801, USA
On Jul 25, 2002 this sequence version replaced gi:46417.
FEATURES
Location/Qualifiers
1..2177
/organism="Rhodobacter sphaeroides"
/mol_type="genomic DNA"
/strain="GA"
/db_xref="taxon:1063"
protein_bind 86..99
/compound="transcriptional activator protein FnrL"
211..1911
CDS
1..2177
/codon_start=1
/transl_table=11
/product="cytochrome oxidase subunit 1"
/protein_id="CAA44514.2"
/db_xref="GI:21955871"
/db_xref="GOA:P33517"
/db_xref="Swiss-Prot:P33517"
/translation="MADAAIHGHDRGRFFRWFNMSNHNKIGVLYLTGSLVGLIS
VAFTVYRMELAPGQVFNCAEHLSEGLVKGFQSLWPSAVENTPNGHLWNVMTGH
```

ORIGIN

```
Query Match 60.0%; Score 1004.8; DB 1; Length 2177;
Best Local Similarity 75.9%; Pred. No. 1.8e-198;
Matches 1294; Conservative 0; Mismatches 377; Indels 33; Gaps 3;

QY 1 ATGCGAGAGCCGCCATTCACGCCCATGACCAACCATGAGCAAGCAAGGCTTCTTTCACGCGC 60
D 1 ATGCGAGAGCCGCCATTCACGCCCATGACCAACCATGAGCAAGCAAGGCTTCTTTCACGCGC 60
Db 211 ATGCGAGAGCCGCCATTCACGCCCATGACCAACCATGAGCAAGCAAGGCTTCTTTCACGCGC 267
QY 61 TGGTTTCATGTCGACCAACCAAGACATCGGTCCTCTATACCTTGTAGCGGTGTGT 120
D 61 TGGTTTCATGTCGACCAACCAAGACATCGGTCCTCTATACCTTGTAGCGGTGTGT 120
Db 268 TGGTTTCATGTCGACCAACCAAGACATCGGTCCTCTATACCTTGTAGCGGTGTGT 327
QY 121 GTTGTTCATGTCGACCAACCAAGACATCGGTCCTCTATACCTTGTAGCGGTGTGT 180
D 121 GTTGTTCATGTCGACCAACCAAGACATCGGTCCTCTATACCTTGTAGCGGTGTGT 180
Db 328 GTTGTTCATGTCGACCAACCAAGACATCGGTCCTCTATACCTTGTAGCGGTGTGT 387
QY 181 GTTGTTCATGTCGACCAACCAAGACATCGGTCCTCTATACCTTGTAGCGGTGTGT 217
D 181 GTTGTTCATGTCGACCAACCAAGACATCGGTCCTCTATACCTTGTAGCGGTGTGT 217
Db 388 GTTGTTCATGTCGACCAACCAAGACATCGGTCCTCTATACCTTGTAGCGGTGTGT 447
QY 218 CGGATGCC-----TCGACAGATGATGAGCAACCAAGACATCGGTCCTCTATACCTTGTAGCGGTGTGT 270
D 218 CGGATGCC-----TCGACAGATGATGAGCAACCAAGACATCGGTCCTCTATACCTTGTAGCGGTGTGT 270
Db 448 CTCTGGCCCTCGGCGGTGAGAAATCGACCCCAAGCAAGGCTTCTTTCACGCGCCTTC 507
QY 271 ACCTACCATGATCTCTGATGATGTTCTTGTGGGTATCCCGCATTTGTTGGGTGTTT 330
D 271 ACCTACCATGATCTCTGATGATGTTCTTGTGGGTATCCCGCATTTGTTGGGTGTTT 330
Db 508 ACCGCGCACCGGATCTCTGATGATGTTCTTGTGGGTATCCCGCATTTGTTGGGTGTTT 567
QY 331 GGTAACATCTGATGCGCTGCAAAATCGGCGCTCCGATATGCGCTTCCCGCGTATGAAC 390
D 331 GGTAACATCTGATGCGCTGCAAAATCGGCGCTCCGATATGCGCTTCCCGCGTATGAAC 390
Db 568 GGCAACTATTTCATGCGCTGCAAAATCGGCGCTCCGATATGCGCTTCCCGCGTATGAAC 627
QY 391 AACCTGCTGTTCTGCGCTGTTTCATGCGCTACCGCATGCGCTTGGCTTGGCTTTCGCGTTCGCA 450
D 391 AACCTGCTGTTCTGCGCTGTTTCATGCGCTACCGCATGCGCTTGGCTTGGCTTTCGCGTTCGCA 450
Db 628 AACCTGCTGTTCTGCGCTGTTTCATGCGCTACCGCATGCGCTTGGCTTGGCTTTCGCGTTCGCA 687
QY 451 CCGGCGCGTGAACGCTGAGTGGGTTCGGGTGTTGGGTTCGTACCGCGCGCTGTCG 510
D 451 CCGGCGCGTGAACGCTGAGTGGGTTCGGGTGTTGGGTTCGTACCGCGCGCTGTCG 510
Db 688 CCGGCGCGTGAACGCTGAGTGGGTTCGGGTGTTGGGTTCGTACCGCGCGCTGTCG 747
QY 511 ACCCGGAGCTGGCTATTCGATGAGCTCGGATTTTCGCGGTTCATTTGCGGTGTC 570
D 511 ACCCGGAGCTGGCTATTCGATGAGCTCGGATTTTCGCGGTTCATTTGCGGTGTC 570
Db 748 ACCTCCGAATCGGCTATTCACCGACCTCGGATCTTTCGCGGTGCACTTGTTCGCGGTGTC 807
QY 571 TCCTTCGATCATGGCGCGATCAACATGATCAGCACTTCTTGAACATGCGCGCGCCCGCGC 630
D 571 TCCTTCGATCATGGCGCGATCAACATGATCAGCACTTCTTGAACATGCGCGCGCCCGCGC 630
Db 808 TCCTTCGATCATGGCGCGATCAACATGATCAGCACTTCTTGAACATGCGCGCGCCCGCGC 867
QY 631 ATGACGCTGCAAAAGTTCGCTTTCGCTGCTGATCTTATCAGCGCTTGGCTGATC 690
D 631 ATGACGCTGCAAAAGTTCGCTTTCGCTGCTGATCTTATCAGCGCTTGGCTGATC 690
Db 868 ATGACGCTGCAAAAGTTCGCTTTCGCTGCTGATCTTATCAGCGCTTGGCTGATC 927
QY 691 CTGCTGGCGCTGCGCGTTCGCTGCTGCAATACCATGCTGTGTCGCGCATCGCACTTC 750
D 691 CTGCTGGCGCTGCGCGTTCGCTGCTGCAATACCATGCTGTGTCGCGCATCGCACTTC 750
Db 928 CTGCTGGCGCTGCGCGTTCGCTGCTGCAATACCATGCTGTGTCGCGCATCGCACTTC 987
QY 751 GGCACGACCTTCTCAATCTCTGCTGGCGCGGTGACCGGATCTTGTACCAACATCTG 810
D 751 GGCACGACCTTCTCAATCTCTGCTGGCGCGGTGACCGGATCTTGTACCAACATCTG 810
Db 988 GGCACGACCTTCTCAATCTCTGCTGGCGCGGTGACCGGATCTTGTACCAACATCTG 1047
QY 811 TGGTCTTCTTGGCGCGCGGATGATCATCATCTTCTGCGCGCTTGGCATCATCAGC 870
D 811 TGGTCTTCTTGGCGCGCGGATGATCATCATCTTCTGCGCGCTTGGCATCATCAGC 870
Db 870 TGGTCTTCTTGGCGCGCGGATGATCATCATCTTCTGCGCGCTTGGCATCATCAGC 870
```

Db 1048 TGGTCTTCGCGCCACCGGAGCTCTACATCATCTGCTGCGCGCTTCGCGATCGTCAGC 1107
Qy 871 CATGTCGTGTGACCTCTCTGAAAGCGCGTCTTCGGTTACCTGCGGATGGTCTATGCA 930
Db 1108 CACGTCATCGGACCTTCGCAAAAGCGGATCTTCGGCTATTTGCGATGGTCTATGCG 1167
Qy 931 ATGTGGCAATCGGTGTCTGCGCTTGTGCTCTGCGCGGACCAACATGATACACGTTGGT 990
Db 1168 ATGTGGCAATCGGCTGCTGCGCTTGTGCTCTGCGCGGACCAACATGATACACGCGGC 1227
Qy 991 ATGTGCGTACCGGACCAATCTCTCTGATGCTGCGGACCAATGATGCGGCTGCGGAC 1050
Db 1228 CTGAGCCTCACCGAGAGCTACTTTATGATGCGGACCAATGATGCGGCTGCGGAC 1287
Qy 1051 GGCATTAAAGATCTCTGCGGATCGGACGATGCGGCGGCTGCGGTTGAGTTCAATCG 1110
Db 1288 GGCATCAAGATCTCTGCGATCGGACCAATGCGGCGGCTGCGGATCGGACCAAGCG 1347
Qy 1111 CCGATGCTGCGGCTTGTGCTTTATGCTCTGCTGCGGCTGCGGCTGCGGCTGCGGATC 1170
Db 1348 CCGATGCTGCGGCTGCGGCTTGTGCTTTATGCTCTGCTGCGGCTGCGGCTGCGGATC 1407
Qy 1171 GTGCTGCGGACCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 1230
Db 1408 GTGCTGCGGACCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 1467
Qy 1231 TTCCATTATGATGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1290
Db 1468 TTCCATTATGATGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 1527
Qy 1291 ATGCGGCAAGTCTGCGGCTGCGGCTTTCGCGGCTGCGGCTGCGGCTGCGGCTGCGG 1350
Db 1528 ATCGGCAAGTCTGCGGCTGCGGCTTTCGCGGCTGCGGCTGCGGCTGCGGCTGCGG 1587
Qy 1351 TTCTTCATCGGTGCGGCTGCGGCTTTCGCGGCTGCGGCTTTCGCGGCTGCGGCTGCGG 1410
Db 1588 ATGCTGCGGCTGCGGCTTTCGCGGCTTTCGCGGCTTTCGCGGCTTTCGCGGCTTTC 1647
Qy 1411 CCGCGGCTTTCGCGGCTTTCGCGGCTTTCGCGGCTTTCGCGGCTTTCGCGGCTTTCG 1470
Db 1648 CCGCGGCTTTCGCGGCTTTCGCGGCTTTCGCGGCTTTCGCGGCTTTCGCGGCTTTC 1707
Qy 1471 GGTGCGTCTGCGGCTTTCGCGGCTTTCGCGGCTTTCGCGGCTTTCGCGGCTTTCG 1530
Db 1708 GCGGCTTTCGCGGCTTTCGCGGCTTTCGCGGCTTTCGCGGCTTTCGCGGCTTTCG 1767
Qy 1531 GTTGCTGCGGCTTTCGCGGCTTTCGCGGCTTTCGCGGCTTTCGCGGCTTTCGCGG 1590
Db 1768 ACGGCGGCGGCTTTCGCGGCTTTCGCGGCTTTCGCGGCTTTCGCGGCTTTCGCGG 1827
Qy 1591 TGGACGCTGCGGCTTTCGCGGCTTTCGCGGCTTTCGCGGCTTTCGCGGCTTTCGCGG 1650
Db 1828 TGGACGCTGCGGCTTTCGCGGCTTTCGCGGCTTTCGCGGCTTTCGCGGCTTTCGCGG 1887
Qy 1651 TGGGCAAGCATCTTCGCGGCTTTCGCGGCTTTCGCGGCTTTCGCGGCTTTCGCGG 1674
Db 1888 TGGGCAAGCATCTTCGCGGCTTTCGCGGCTTTCGCGGCTTTCGCGGCTTTCGCGG 1911

RESULT 5
LOCUS PDCOX2 3211 bp DNA linear BCT 07-OCT-1996
DEFINITION P.denitrificans COI gene.
ACCESSION X05829
VERSION X05829.1 GI:994806
KEYWORDS COI gene; cytochrome c oxidase; cytochrome oxidase a3;
unidentified reading frame.
SOURCE Paracoccus denitrificans
ORGANISM Paracoccus denitrificans
Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
Rhodobacteraceae; Paracoccus.
REFERENCE 1
AUTHORS Raitio,M., Jalli,T. and Saraste,M.

TITLE Isolation and of the genes for cytochrome c oxidase in Paracoccus denitrificans
JOURNAL EMBO J. 6, 2825-2833 (1987)
REFERENCE 2
AUTHORS Saraste,M.
TITLE Direct Submission
JOURNAL Submitted (04-JAN-1991) Saraste M., EMBL, Meyerhofstrasse 1, Heidelberg, Germany
REMARK revised by [3] MAT
REFERENCE 3 (bases 1 to 3211)
AUTHORS Saraste,M.
TITLE Direct Submission
JOURNAL Submitted (20-SEP-1995) Saraste M., EMBL, Meyerhofstrasse 1, Heidelberg, Germany
COMMENT On Sep 21, 1995 this sequence version replaced gi:45474.
FEATURES
source
1..3211
/organism="Paracoccus denitrificans"
/mol_type="genomic DNA"
/db_xref="taxon:266"
/map="cox2 locus"
218..222
RBS
229..564
CDS
/note="unnamed protein product; ORF4"
/codon_start=1
/transl_table=11
/protein_id="CAA29273.1"
/db_xref="GI:45475"
/db_xref="GOA:P08304"
/db_xref="Swiss-Prot:P08304"
/translation="MADISMPKASALGETSRAGASIDDAFLALSKVVFADGALDKRT
KQLIAYAVAVHTQCPWCIEGHVKGARREGATNEQIMEAIWVAEMRAGAAAYAHAKAL
AVLDEIDIGA"
229..561
mat_peptide
/product="ORF4"
628..640
repeat_unit
/note="inverted repeat A"
686..689
TATA_signal
713..725
repeat_unit
/note="inverted repeat A"
765..769
RBS
777..2444
gene
/gene="COI"
777..2444
mat_peptide
/gene="COI"
/product="cytochrome-c oxidase"
/EC_number="1.9.3.1"
777..2441
CDS
/gene="COI"
/EC_number="1.9.3.1"
/codon_start=1
/transl_table=11
/product="cytochrome-c oxidase"
/protein_id="CAA29274.1"
/db_xref="GI:995678"
/db_xref="GOA:P08305"
/db_xref="Swiss-Prot:P08305"
/translations="MSAQISDSIEKKRGFTFRMFSTNHKIDGLVLYFTAGLAGLTSV
TLTVYMELOHPGVOYMCLEGLMVLADAAECTPAHLNVTYVTVYGLIMPFVVP
ALFGGCTFMPFHLIGAPDPAFRLNLSYWLXVCGVSLASLSLSPGSDQPCAGV
WVLYPSTTEAGAMDALFAVHSGATSGILGAINITFLNMRAPGNTLTKVPLFA
WAYFIATMILSLVLAGGITMLMDRNFQFFDPAGGDPVLYQHILWFFGHEV
YMLIFPGIISHVISTFARKPIFGYLPMLAAIAFLGFIWAHMYTAGMSLQQ
TFQMATMTIATVPTGKVFWSIATMGGSEIEFKTLMWALFLFTVGGTGVVIAQGS
LDRVYHTVYVAHPTVMSLGAIPAI FAGTVYWGKMSGROYPEWAGOLHFMFMFIG
SNLIFEPQFLPGQMPRVIDYVPEFSYNNISIGAVISFASLFFIGTVFYTLFA
GKPVNPNYNNHADTLEWLPSPPEHFEILPDPEDWDRAQHR"
2744..2788
terminator
2744..2755
repeat_unit
/note="imperfect inverted repeat B"
2764..2775
repeat_unit
/note="imperfect inverted repeat B"

ORIGIN

QY	1031	TGGTGATCGCGGTGCGACCGGCATTAAGCATCTCTCGTGGATCGCCAGATGTGGGGCG	1090
Db	1801	TGACCATCGCGGTGCGACCGGCATCAAGGTCTTCTGATGATGCGACCATGTGGGGCG	1860
QY	1091	GCTCGCTTGAGTTCAAATCGCGATGCTCTGGGCGCTTTTGGCTTTATGTTCTCTTCAACG	1150
Db	1861	GCTCGATCGAGTTCAAGACGCGATGCTCTGGGCGCTTG	1914
QY	1151	TGGGTGGTGACCGGTATTCGTCTGGCCCAAGCGGCTTGACCGTGCATATCAGACA	1210
Db	1915	TGCGGCGGTGACCGGGGTGTCATCGCCACAGGGTCCCTTGACCGCGTTATCAGACA	1974
QY	1211	CCATTACGTGGTGGCGCACTTCCCATATGATGATGCTGGGTGGATGGAATTTGGCATCT	1270
Db	1975	CCATCTATATCGTGGCGCATTTCCATATGATGAGCCCTTGGGCGCTCTTCGCGATCT	2034
QY	1271	TCGCGGTATCTACTTTTTACATCGCGAAGTTCTGGGCGCGGCTTTCGCGGAATGGCTG	1330
Db	2035	TTGAGGACACTATTACTGGATCGGCAAGATGTCGGGCGGCAATACCGGGAATGGCGG	2099
QY	1331	CAAGCTGCACCTTCGACCTTCTTCATCGGTGCGAAGCTGACGTTCTTCCCGCAGCACT	1390
Db	2095	GCCAGCTGCAATTTCTGGATGATGTTTATCGGCTCGAACCTGATCTTCTCCGCGACACT	2154
QY	1391	TCCTGGGACGTCAGGGTATGCGCGCGCTTACATCGACTATCCGGAAGCCTTCGCGCTGT	1450
Db	2155	TCCTGGCGCCAGGCGCATCGCGCGCGCTATATCGACTACCGGTCGAGTTCTCGTACT	2214
QY	1451	GGAAACAAGTCTCGTCTATGTTGGTTCGCTCGGCTTCGCTCGTCTTCTTTCATCG	1510
Db	2215	GGAAACAACATCTCGTGGATCGGCGCGCTATATCTCTTGGCTCTTCTTCTTTCATCG	2274
QY	1511	TGATCTTTGTCTATAGCTGTGCTGGCGCGCGCGAGACCCGTCGGAACCCGTGGGGG	1570
Db	2275	GCATCGTGTCTACACGCTCTTCGCGGCAAGCCCGTGAACGTGCGCACTTCTGAAACG	2334
QY	1571	AATTGCGCGATAGCTGGATGAGCGTGCATATCAGGCTCTCGGCGCCACACCTTCGAAA	1630
Db	2335	AGCATGCGGACACGCTGGAATGGAACCGCTCGCCCTCGCGCGCGGAGCATCTTCGAGA	2394
QY	1631	CGCTGCCCAAGCGCTCGGACTGGGA	1655
Db	2395	CCCTGCCCAAGCTGAGACTGGGA	2419
RESULT 6			
SME591785		286550 bp	DNA linear BCT 05-JUL-2002
LOCUS			
DEFINITION			Sinorhizobium meliloti 1021 complete chromosome, segment 4/12.
ACCESSION			AL591785 AL591688
VERSION			AL591785.1 GI:15073719
KEYWORDS			
SOURCE			Sinorhizobium meliloti (Rhizobium meliloti)
ORGANISM			Sinorhizobium meliloti
REFERENCE			Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
AUTHORS			Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
			1 (bases 1 to 286550)
			Capela,D., Barloy-Hubier,F., Gouzy,J., Bothe,G., Ampe,F., Batut,J.,
			Boisard,P., Becker,A., Boutry,M., Cadieu,E., Dreano,S., Gloux,S.,
			Godrie,T., Goffeau,A., Kahn,D., Kiss,E., Lelaure,V., Masuy,D.,
			Pohl,T., Portetelle,D., Puehler,A., Purnelle,B., Rameperger,U.,
			Renard,C., Thebault,P., Vandenbol,M., Weidner,S. and Galibert,P.
			Analysis of the chromosome sequence of the legume symbiont
			Sinorhizobium meliloti strain 1021
			Proc. Natl. Acad. Sci. U.S.A. 98 (17), 9877-9882 (2001)
JOURNAL			
MEDLINE			21396507
PUBMED			11481430
REFERENCE			2 (bases 1 to 286550)
AUTHORS			Gouzy,J.
TITLE			Direct Submission
JOURNAL			Submitted (26-JUL-2001) Gouzy J., Submitted on behalf of the MELIOL
			EU Consortium
COMMENT			MELIOL EU Consortium:

Laboratoire de Biologie Moléculaire des Relations
Plantes-Microorganismes, UMR215-CNRS-INRA, BP27, F-31326 Castanet,
France, Laboratoire de Génétique et Développement UMR6061-CNRS,
Faculté de Médecine, 2 avenue du Pr. Léon Bernard, F-35043 Rennes,
France, GATC GmbH, Fritz-Arnold-str. 23, D-78467 Konstanz, Germany,
Universitaet Bielefeld, Biologie IV (Genetik) Universitaetstr. 25,
D-33615 Bielefeld, Germany, Unite de Biochimie physiologique,
Université Catholique de Louvain, Place Croix du Sud 2, Bte 20,
B-1348 Louvain-la-Neuve, Belgium, Unite de Microbiologie, Faculté
des Sciences Agronomiques de Gembloux, Avenue Marechal Juin 6,
B-5030 Gembloux, Belgium. E-mail:Jerome.Gouzy@tcoulouse.inra.fr
<http://sequence.toulouse.inra.fr/meliloti.html>.

FEATURES

```

source
1..286550
/organism="Sinorhizobium meliloti"
/mol_type="genomic DNA"
/strain="1021"
/db_xref="taxon:382"
70..813
/gene="gst11 OR SMC00916"
70..813
/gene="gst11 OR SMC00916"
/EC_number="2.5.1.18"
/function="miscellaneous"
/notes="Product confidence : putative
Gene name confidence : putative
predicted by Codon usage
predicted by Homology
predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PUTATIVE GLUTATHIONE S-TRANSFERASE PROTEIN"
/protein_id="CAC45361.1"
/db_xref="GI:15073720"
/db_xref="GOA:Q92RR5"
/db_xref="TrEMBL:Q92RR5"
/translations="MADLSFFPIIARWPAKNPDIILQSLTPNGVKVSAIEELGLA
YEAFISFDNEQSPFSLNPNRPAIDPNPDGKPIGLFESGAILIYLAETKG
KLIPADAAEYELCYWFMCGFMFGFHFPAADKVANNPVPVRYDEAKR
LLGVLEARLEGROWLMGRNTPTSPTRPIMWKGKSSMVAPRCSTIRASMSWTGI
AAWHALRPREARSRARSRSRSLPAGD"
588..779
/gene="SMC00915"
588..779
/gene="SMC00915"
/functions="miscellaneous; hypothetical/partial homology"
/notes="Product confidence : hypothetical
Gene name confidence : hypothetical
predicted by Codon usage
predicted by Homology
predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="HYPOTHETICAL PROTEIN"
/protein_id="CAC45362.1"
/db_xref="GI:15073721"
/db_xref="TrEMBL:Q92KL5"
/translations="MADGGEEYTIADIATYPWVEGARPYGGAELVDYKSPFNVDWV
DRGLAPAAQKGMIEPRKE"
867..2084
/gene="SMC00914"
867..2084
/gene="SMC00914"
/function="small molecule metabolism"
/notes="Product confidence : putative
Gene name confidence : hypothetical
predicted by Codon usage
predicted by Homology
predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="HYPOTHETICAL PROTEIN"
/protein_id="CAC45363.1"
/db_xref="GI:15073722"
/db_xref="GOA:Q92RR4"
/db_xref="TrEMBL:Q92RR4"
/translations="MSGRLLVVGGQAQAFALVAKLRKONRPTTVVAERASLPYQRP
PLSKYLLREMTLDRLLYRPEAWAHEIDIRLSTTVTRDLAKQVALSDGSMITVE
TLAFTATGTPRLPAAVGGDLAGVVFVRDFRADRLAEMQGRVILVVGGLYGLGA
AAVATSGLEVTIEMADRIQVVASATSAIVEIRHSHGVDIRERTGLHRLICDNG
RVTAELSDGSGVIPVDIVIVIGVAANDALAEAGIETANGIVVDVSHGRSDPTIVAM
GCAVLPMWGMRIRESVQNAVDQAEAAVAAGLAGTDPDYPKPMFSDQYDKLIQAG
FGLGHDTLVFGQSGSVVYFRQGLIANDAINAKAVVTKKLEAGATPDRLL
LADQVDLAKALLV"
Complement (2114..2572)
/notes="Sm-2 OR SMC04876
REPEAT SM-2
Predicted by Homology"
/evidence=not_experimental
/complement (2585..4222)
/gene="groEL1 OR SMC00913"
/complement (2585..4222)
/gene="groEL1 OR SMC00913"
/function="cell processes; chaperoning"
/notes="predicted by Codon usage
predicted by Homology
predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="60 KD CHAPERONIN A PROTEIN"
/protein_id="CAC45364.1"
/db_xref="GI:15073723"
/db_xref="GOA:P35469"
/db_xref="Swiss-Prot:P35469"
/translations="MAAKEVFGRSAREKMLRGVDILADAVKVTLPKGRNVVIDKSF
GAPRTKDGVSVAKEIELEDFENMGQMYREASKNTDIAGDGTATVLAQAIARE
GAKAAVMGPMDLKRGIDILAAVEVXKDLAKAKINTIDEVAQVGTISANGKQIGL
DIAMAMQNGEVITVEEAKTAELEVEVQKQFDRGYSYPVTNPKRWADLEDA
FILLHKLNLQALPVLGAVVCTGKPLIIAEDVEGEALATLVNKLKGGKIAAV
KAPGFRRRKMLIEDAILTGGTVISDLGKLESVTLMLGRKAKVSIKTENITVD
CAGKSDIERVAQIKAEIETISDYSIDREKQLAKLAGEAVAVRGGATEVEVREK
KDRIDALNATRAAVQEGIVFGGVALLRSSVKITVGENDDQDQAGVNIVRRLQSPA
KQIVENAGDEASIVVQKIDLEKNTDDFGYNAQTGEYGDMIAMGIIDPVKVVRTALQDAA
SVASLITTEAMTAELPKDAPAMPGGMGGMGMW"
Complement (4298..4594)
/gene="groES1 OR SMC00912"
/complement (4298..4594)
/gene="groES1 OR SMC00912"
/function="cell processes; chaperoning"
/notes="predicted by Codon usage
predicted by Homology
predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="10 KD CHAPERONIN A PROTEIN"
/protein_id="CAC45365.1"
/db_xref="GI:15073724"
/db_xref="GOA:P35473"
/db_xref="Swiss-Prot:P35473"
/translations="MASTNFRPLHDRVVRVSEKTKGIIIPDTAKEKPOGSEIV
AVGSGARDESGKVPLDVKAGDRILFKKSGTEVKINGEDLLIMKEADINGVIG"
4863..5690
/gene="SMC00911"
4863..5690
/gene="SMC00911"
/functions="miscellaneous; hypothetical/global homology"
/notes="Product confidence : hypothetical
Gene name confidence : hypothetical
predicted by Codon usage
predicted by Framed"
/codon_start=1

```



```

RESULT 7
LOCUS AE009044
DEFINITION Agrobacterium tumefaciens str. C58 linear BCT 28-MAY-2004
of 256 of the complete sequence.
ACCESSION AE009044 AE008688
VERSION 1
KEYWORDS Agrobacterium tumefaciens str. C58
SOURCE Agrobacterium tumefaciens str. C58
ORGANISM Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
REFERENCE
AUTHORS Wood D.W., Setubal J.C., Kaul, R., Monks, D., Chen, L., Wood, G.E.,
Chen, Y., Woo, L., Kitajima, J.P., Okura, V.K., Almeida Jr., N.F.,
Zhou, Y., Bovee Sr., D., Chapman, P., Clendenning, J., Deatherage, G.,
Gillet, W., Grant, C., Guenther, D., Kutyavin, T., Levy, R., Li, M.,
McClelland, E., Palmieri, A., Raymond, C., Rouse, G.,
Saenphimmachak, C., Wu, Z., Gordon, D., Eisen, J.A., Paulsen, I.,
Karp, P., Romero, P., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M.,
Krespan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C.,
Zhao, Z., Dolan, M., Tingey, S.V., Tomb, J., Gordon, M.P., Olson, M.V.
and Nester, E.W.
The genome of the natural genetic engineer Agrobacterium
tumefaciens C58
Science 294 (5550), 2317-2323 (2001)
21608550
PUBMED 11743193
REFERENCE
AUTHORS 2 (bases 1 to 10790)
Wood D.W., Setubal J.C., Kaul, R., Monks, D., Chen, L., Wood, G.E.,
Chen, Y., Woo, L., Kitajima, J.P., Okura, V.K., Almeida Jr., N.F.,
Zhou, Y., Bovee Sr., D., Chapman, P., Clendenning, J., Deatherage, G.,
Gillet, W., Grant, C., Guenther, D., Kutyavin, T., Levy, R., Li, M.,
McClelland, E., Palmieri, A., Raymond, C., Rouse, G.,
Saenphimmachak, C., Wu, Z., Gordon, D., Eisen, J.A., Paulsen, I.,
Karp, P., Romero, P., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M.,
Krespan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C.,
Zhao, Z., Dolan, M., Tingey, S.V., Tomb, J., Gordon, M.P., Olson, M.V.
and Nester, E.W.
Direct Submission
Submitted (27-SEP-2001) Department of Microbiology, University of
Washington, 1959 NE Pacific Ave, Box 357242, Seattle, WA
98195-7242, USA
FEATURES
source
Location/Qualifiers
1..10790
/organism="Agrobacterium tumefaciens str. C58"
/mol_type="genomic DNA"
/strain="C58"
/isolate="U. Washington"
/db_xref="taxon:176299"
complement(72..596)
/genes="ialB"
/notes="synonym: Atu0766"
complement(72..596)
/genes="ialB"
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="invasion associated locus B"
/protein_id="AAL41782.1"
/db_xref="GI:17739136"
/translacion="MELSPALFTFLAAGLAIAPAAALAAQQPPGTPKSTHGAWSV
ICKPAGASEECALMQVIADREVLGSLVILTAADKRIKILAPLGLVLDKGM
EYIINDNIGRAIFRCFSEGCYVEVDIDDELKVLRAKNVAFALRESVDQDRGIP
IISLFGAEGYDALP"
1029..1892
/genes="coxB"
/notes="synonym: Atu0767"
1029..1892
/genes="coxB"
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"

```

```

/codon_start=1
/transl_table=11
/product="cytochrome c oxidase subunit II"
/protein_id="AAL41783.1"
/db_xref="GI:17739137"
/translacion="MTCLLTAVGAHADQPVHQMGMQBAATPIMHEIRWEQYTLWFI
VFTLVLLALLIIVAVKPHAKNPVASKTSHTAIEVWTALPVLILLLFAFSPNLL
NAOLTPOENPDLTILKATATOWLSYKYAAEAPLPSFDSYLLKQDRAAAGKEDKAR
YPRLLADNEMVVPVGTVRLVLTAAPTDVIHAPAMPAGVKIDKAPCRNLNETWFKPE
KEGLYGCSELCKGDHAFMEPIARVVEQQYNTWHAASADLNGANRALMASVDGAP
RTVDAANETN"
1901..3619
/genes="coxA"
/notes="synonym: Atu0768"
1901..3619
/genes="coxA"
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="cytochrome-c oxidase chain I"
/protein_id="AAL41784.1"
/db_xref="GI:17739138"
/translacion="MTTMAGPSAHDDHSHGQSAHDAHADHSHDHSKPGPFARF
LSTNHKDIGTLYLIFALMAGSIIGGLSVNRMELQEPGIOIPHCLASMYVGFSDAI
DGGKHENFVTTAHALINIFMWPAMPAMIGGFANWIFIMIGADPMAFPRLNNISFWLI
VPFILLLSLUFVEGPGAGYGGGWTMPPLSTSGMPGPAVDLAIPLSHVAGASSIL
GAINFITILLNRAPGMTLHKMPLFAMSVLPTAFLLLSLPVLGAGITMLLTDNFEQT
AFSPGEGDPIYQHLFWFPGHPEVYLLPFGFIISHIHSVTFSKPKVFGYGLMAYA
MVAIGAGFVIVAHMVTGLSLEAQRVFPVATWIAVPTGKIFISWIAVMGCSLTF
STWVWAIGFIFLTGVGTVOLANAGLDRSLDHYVVAHPHYVLSLCAVPAIAG
MYWFPKITGYMNERIKGLHFWMPITGVNLIFFPOHFLGLACMPRYIDYDAYAGM
NMSSVGSYISAVAGTFLFGVNEAPAKRIAGNFWEGATTLEQLSPPPPHOWE
QLPAPR"
4009..4962
/genes="ctaB"
/notes="synonym: Atu0769"
4009..4962
/genes="ctaB"
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="protoheme IX farnesyltransferase"
/protein_id="AAL41785.1"
/db_xref="GI:17739139"
/translacion="MTVIDDRDMGAESELSSEAGARDYFELLKPRVMSLVVFTAPAG
LVLAPEINFLGLIALLCIAVGAGASGALNMWDADIDAVMTIRAKRIPSGRIAPR
EALAFGLTSAFSAVILGUAVNWSAGLIAFTIFFYAVYVYTWLKESTPQNTIVIGAA
GAPFPMGLNACVTGVSLSVILFLIIFLWPAHFWALFKMRDYGAVGIPMNMVA
GERSTNQMIYAVLTAAAVAPFTGLASAGYGFIAVLISALFVYCSFVDRMPEGD
EKMLPAKQNPAYSVLVLFALFISGLLADHAPALKAIVISGLV"
5005..5739
/genes="coxG"
/notes="synonym: Atu0770"
5005..5739
/genes="coxG"
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="cytochrome C oxidase assembly protein"
/protein_id="AAL41786.1"
/db_xref="GI:17739140"
/translacion="MGAILHSASCSPPVWCCSTSPSSRSARIDGLS:IGGEDMANME
AETARAGSORVNRSLVLCVFFCAIMGMAVAPVLSLFCRVTGNGTGTVOYQS
DVLIDKTIIVTFDANTSGLNMDFRVDKMPKIGETIKVTFKATNSPVATGTAV
ENVTMEAGAYFNKVECFCTETTLQBGETLEMPVVFVDPDIAIARETKSVHTILS
YTFYPAKTEKPVASLPVKTESGDSKL"
5801..6676
/genes="coxC"
/notes="synonym: Atu0771"
5801..6676

```


/gene="coxC"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="cytochrome c oxidase subunit III"
/protein_id="AAL41787.1"
/db_xref="GI:17739141"
/translation="MADTHQKHNDYHIDPSPPLASIGAFIMTFGVCVMRYLSGG
SPKFGAELANPMLPYGLIVLVYVMWADTIKEANEGSHTRVLSHLRYGMIMFI
ASEVNFVAMFWAFDASLPFHEAIQASRLLEYTGQMPKGIIEVDPMHPLXNTVIL
LLSGTCVWMAHLLHNDKGLISGLALTVALGSLVQVYEVIHAFPFKRSIYGA
TFFNATGPHGFHVFVGTVELLVCLFRAIAGGFTPKQHFGEFAAAWYHWFVDVWLFILF
FAIYIWWGAGAPLHG"
6938..7324
/gene="Atu0772"
6938..7324
/gene="Atu0772"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAL41788.1"
/db_xref="GI:17739142"
/translation="MSEPSNGNFAPVDPIKVLKGCPCRCNGRLFDGFLTPKACSA
CGLDYFADAGDGPVFMVILVFGVLMALVDORFAPFVMAHMLWLPFTVIVSLV
LLRKLKGMIALQYRNNASEGRLDRE"
7293..8069
/gene="surfl"
/note="synonym: Atu0773"
7293..8069
/gene="surfl"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="surfeit 1"
/protein_id="AAL41789.1"
/db_xref="GI:17739143"
/translation="WQAKGLIVNDVSHKAAQRVWFAAPLVLLALAILGLTWQV
KLYWEALMADIERNAPSLDIEAIVKSGEIEYKRVLSGTFDTHREHFFA
THQGTGYIYVTLADQRIPLFVNRGFVPEFMEKAARPDQGVSEVITIGLARAPL
VAKPSLLPDNDIAKNIYWKDLAAMASSAEVPERLVKLFVDADNAPNPGWPQGV
TLIDLPNHLQVATWYGLAALVIVAGFAYFRNGKAOGE"
8128..9174
/gene="lytB"
/note="synonym: Atu0774"
8128..9174
/gene="lytB"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="penicillin tolerance protein"
/protein_id="AAL41790.1"
/gene="coxC"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="cytochrome c oxidase subunit III"
/protein_id="AAL41791.1"
/db_xref="GI:17739144"
/translation="MADTHQKHNDYHIDPSPPLASIGAFIMTFGVCVMRYLSGG
SPKFGAELANPMLPYGLIVLVYVMWADTIKEANEGSHTRVLSHLRYGMIMFI
ASEVNFVAMFWAFDASLPFHEAIQASRLLEYTGQMPKGIIEVDPMHPLXNTVIL
LLSGTCVWMAHLLHNDKGLISGLALTVALGSLVQVYEVIHAFPFKRSIYGA
TFFNATGPHGFHVFVGTVELLVCLFRAIAGGFTPKQHFGEFAAAWYHWFVDVWLFILF
FAIYIWWGAGAPLHG"
6938..7324
/gene="Atu0772"
6938..7324
/gene="Atu0772"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAL41788.1"
/db_xref="GI:17739142"
/translation="MSEPSNGNFAPVDPIKVLKGCPCRCNGRLFDGFLTPKACSA
CGLDYFADAGDGPVFMVILVFGVLMALVDORFAPFVMAHMLWLPFTVIVSLV
LLRKLKGMIALQYRNNASEGRLDRE"
7293..8069
/gene="surfl"
/note="synonym: Atu0773"
7293..8069
/gene="surfl"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="surfeit 1"
/protein_id="AAL41789.1"
/db_xref="GI:17739143"
/translation="WQAKGLIVNDVSHKAAQRVWFAAPLVLLALAILGLTWQV
KLYWEALMADIERNAPSLDIEAIVKSGEIEYKRVLSGTFDTHREHFFA
THQGTGYIYVTLADQRIPLFVNRGFVPEFMEKAARPDQGVSEVITIGLARAPL
VAKPSLLPDNDIAKNIYWKDLAAMASSAEVPERLVKLFVDADNAPNPGWPQGV
TLIDLPNHLQVATWYGLAALVIVAGFAYFRNGKAOGE"
8128..9174
/gene="lytB"
/note="synonym: Atu0774"
8128..9174
/gene="lytB"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="penicillin tolerance protein"
/protein_id="AAL41790.1"

Query Match 44.9%; Score 750.8; DB 1; Length 10790;
Best Local Similarity 68.0%; Pred. No. 1e-145;
Matches 1109; Conservative 0; Mismatches 512; Indels 11; Gaps 4;
13 GCCATTACAGCCGATGACCAACCATGAGAGCAAGGCTTCTTACGCGCTGGTTCATGTCG 72
1979 GACGATCATTCGATGATCATCCCATAGCCGGGCTTTTCGCCGCTGGTTCCTGTCG 2038
73 ACCAACCAAGACATCGGTTCGTATACCTTGAGCGCTGGTGTGTTGGTTTCATT 132
2039 ACCAACCAAGACATCGACGCTGTACCTGATCTTCGCCATCATGCGGGGTATCAIC 2098
133 TCCGTCTGTTTCCCGCTCTACATCGGCTTGGAGTGTGATGATCGGGGTGTTCAGTACATG 192
2099 GCGCGCGGGCTTTCGGTCTGTCATCGGTATGAGCTGACGAGCGCGGCATCCAGATCTTC 2158

193 TGCTTTGAAGGGCGCAGCTCTGATCG-----CGATGCTCGCAGACATGTACGGCGAACGG 248
2159 CACGGTCTGGCTTCCATGCTTACGGCTTCGAGGGCGATGCCGCCATCGACGGCGGTAAAG 2218
249 ACACCTGTGGAACGTCATGTTTACCTACCATGGTATTCTGATGATGATTTCTTTGTGGGTAT 308
2219 -CACATGTTCAAGCTGTTACGACGCGACGCGCTGATCATGATCTTCTCATGCTCAT 2277
309 CCCCGCATTTTCGGTGGTGGTAACTATCTGATGCGCGCTGCAAAATCGGCGCTCGGA 368
2278 GCGCGCATGATCGCGGGTTCGCCAACTGGATGATCCCGATCATGATCGGCGCGCTGA 2337
369 TATGGCTTCCCGGCTATGAACAACCTGCTGCTTCTGCTGTTCAATGCGCGGTACCGCGAT 428
2338 CATGGCTTTCGCCGGCTCAACAATTAATTCCTTCTGCTCACTGTTCCGGCTTCATCT 2397
429 GGGCGTGGCTTCTGCTTTCGACCGCGCGGTGACGCTGAGCTGGTTCGGGCGCTTGGTTG 488
2398 GCTGCTCTCTGCTCTTCTGTCGAGGTCGCGGGGTGCTTATGGCGTGGGTGGCGCTG 2457
489 GGTTCGTGTACCGCGCGCTGTCGACC---CGCGAAGCTGGCTATTCGATGACCTCGCGAT 545
2458 GACCATGATTCGCGCTCTGTCGACCACTGTCGCGCGCGCGGTGGATCTTGGGAT 2517
546 TTTCCGGTTCACCTTGTGCGGTGCTCTCTGATCATGCGCGCGATCAACATGATACGAC 605
2518 CTTCGCTGCGATCGCGCGGTGATCTCTGATCTCTGCGCGCATCAACTTCATCACCA 2577
606 CTTCCTGAAATGCGCGCGCGCGCGGTGACGCTGCAAAAGTGGCTTCTTCTGCTGTC 665
2578 CATCTGAACATGCGCGCGCGCGCATGACGCTGCAAAAGTGGCTTCTTCTGCTGTC 2637
666 GATCTTATACGCGCTTGGCTGATCTCTGCTGGCGCTGCGCGTCTGCTGCTGCTGCAATCAC 725
2638 GGTTCGTGTACCGCATCTGCTCTCTGCTTCTGCTTCTGCTTCTGCTGCTGCTGCTG 2697
726 CATGCTCTGACCGACCGTAACTTTCGCGACGACCTTCTTCAATCTCTGCTGCGCGGTG 785
2698 CATGCTTCTGACCGACCGTAACTTTCGCGACGCGCTTCTTCTGCGCGGAGCGCGGTG 2757
786 CCCGATCTGTACCAACACATCTCTGCTGCTTCTTGGCACCGCGAAGTGTACATCATCAT 845
2758 TCCGATCTCTACCGACGATCTGTTCTGCTTCTGCTGCTCAACCGGAAAGTGTACATCTCAT 2817
846 TCTCCCGCGCTTGGCATCATGACCCATGTCGTCGACCTTCTGCAAAAGCGCGCTCT 905
2818 CTGCGCGGTTCGGCATCATGACCCACATCTGTCGACCTTCTTCCAAAGACCGCGCTCT 2877
906 CGGTTCCTGCGGATGTCATGCAATGGTGGCAATCGGTGTTCTGCGGCTTGTGCTGCTG 965
2878 CGGTATCTCGGCGATGGCTTACGCGCATGGTGGCATGGTGGCTTCTGCTGCTGCTGCTG 2937
966 GCGCACCATGATACACCGTGGTATGTCGCTGACCCGACCAATCTTACTTTCATGCTGGC 1025
2938 GCGCACCATGATGATACCGTGGTATGTCGCTGACCCGCTGCTGCTGAGGACAGCGCTACT 2997
1026 CACCATGATGATGCGGCTGCGACCGGATTAAGATCTTCTGCTGGATGCGCAGATGTCG 1085
2998 CACCATGATGATGCGGCTGCGACCGGATTAAGATCTTCTGCTGGATGCGCAGATGTCG 3057
1086 GGGCGGCTCGGTGAGTTCAAAATCGCGATGCTCTGCGGCTTCTGCTTATGTTCTCTGCT 1145
3058 GGGCGGCTCGGTGAGTTCTCGACCCGATGGTCTGCGCATCGGCTTATCTTCTCTGCT 3117
1146 CACCGTGGGTGGTGGACCGGTATGTCGTCGCGCAAGCGGTGTGGACCGGTGATATCA 1205
3118 CACCGTGGGTGGTGGACCGGTATGTCGTCGCGCAAGCGGTGTGGACCGGTGATATCA 3177
1206 CGACACCTTATACGTGGTGGCGCATCTCCATTTGATGTCGCTGGGTGGATCTTTCG 1265
3178 CGACACCTTATACGTGGTGGCGCATCTCCATTTGATGTCGCTGGGTGGATCTTTCG 3237
1266 GATCTTCGCGGTATCTACTTTTACATGCCGAAGTTCTCGGCGCGCGCTTTTCCCGGAATG 1325

VOGAEKIVVSGPKSVYDRLKVEGGRISLIDINLADGAIMTLGPDGLQVRSDRDMKI
TVVAPDVTREVEGSGDIDIRRYEQALNLDVSGDVEAGKVTETVTVENSNGSYAD
LGLKAEAALEIVSGSGAVPATGKAKIDISSGDVELRTEPKSVTSEITGSGEVR
EY"

gene

4964..6436

CDS

/gene="CC3411"

4964..6436

/note="CC3411"

/note="identified by match to PFAM protein family HMM

PF01523"

/codon_start=1

/transl_table=11

/product="13D protein"

/protein_id="AAK25373.1"

/db_xref="GI:13425123"

/translation="MSAFGLTCRTNAPAPIPASSILIDAAAGVDPDRARAILGSPALA

RADDGLFLERSSEAFVDDGRLKASYSDSGEFLRVAGETAGYAHAAEISPAAI
RRASDSALAKRHAGVABGPATNEXLYGENDPVSAFDFDKVALLQEIADAFARAR
DRPVOMASWAGERVWEILRADGLRIDVRPLRVNVQVTVKDGRRSGSGAGG
RAGFPAWISPKWDQDOVDLALMALNLDVACPAAGMDVVLGPGWNVLHEAVGCH
LEGDFNRKGISASGSGIGERVAAPCVTVFDDSGISPGRRSLTVDEGPTERTILIED
GLIVGYNHDSARLMGMKATGNRRQSTAHMPMPMTNIGLAGEHDPQEMIASLKR
NDFDFDGIQVCGKSGGVPGVGVGPSLKITGLTVGGTAV"

gene

/gene="CC3412"

CDS

complement(6601..7098)

/note="CC3412"

/note="identified by match to protein family HMM"

/codon_start=1

/transl_table=11

/product="acetyltransferase, GNAT family"

/protein_id="AAK25374.1"

/db_xref="GI:13425124"

/translation="MSRLSIPILADLNADPDGAALLNNAHAVELSMLEPERLAHV
AQAFARRVADALLTDEHDADYSPNPLFRERYPPVYVDRVAVDASARGELA
RALYDDLDAKAAAGHARIVCEVNSDPNPVSDAFHAALGVFVGTAEIHGKNTVY
LERRL"

gene

complement(7386..9266)

CDS

/gene="CC3413"

complement(7386..9266)

/note="CC3413"

/note="identified by match to protein family HMM"

/codon_start=1

/transl_table=11

/product="peptide transporter"

/protein_id="AAK25375.1"

/db_xref="GI:13425125"

/translation="MNIVAVGILVTLVTGIPVLLQLLKHPRGLIICFLABMWRFS
YVGMRAILFYLTQHLFDPKASGHVSTALVYLVLIGGLADKMLGTRKAVAPG
AVLLVAGHLAMATEGEAVQCTLYQCATVEFOAEGRGONEAKLVUDGKAVEATKV
GDEIKGLPANAPLPAVLKADYKLDVAGRDLYLNIFWFSLSLILGFLKPNIST
IVQGLPQGPFRDPGFTLYIYGINUGSPWASILCGLLVNVVWAGFLAGVGLAG
FVFLGKPMDDKAPPEPKVLEKRVAGPINRELAISLVGLVGLVFFLVQFNAY
GVANTGILASLGYIGFFMFTKCAKEERERLALAVELIFGAVFWTLFEQAGSLSLF
AANTVQDLVREPLRPLFSGAVTLATPDQAKAGIDAASTFWVDTSFNAPQTCALNAGW
ILIFAPVFAVMTFLGAGKDPNPVKFGLALIQVAGFLQLQWGSTFADGAPKPLI
FLVMVLMHTTGELCKSPVGLSOMTKLSPASVVSFMMVAWFLAVIAIAOWVGSKLAGLA
AETVGGQVLDPGMALAASLKVENIIGLVAVALIGIGFLLLSFVLKXWNSHADNVGAPK
XA"

gene

complement(9726..11087)

CDS

/gene="CC3414"

complement(9726..11087)

/note="CC3414"

/note="similar to GP:3688156; identified by sequence

similarity; putative"

/codon_start=1

/transl_table=11

/product="glutamate-cysteine ligase"

/protein_id="AAK25376.1"

/db_xref="GI:13425126"

/translation="MADTASVQERQLTDLDTAYFAKSGKPEAFRVGAHEKFGFYL
GSEHPVYEGDKGVHALLTCLQRFQWTPVMEGPTIIGLRNGANVSLRPGQFELSGA

gene
5167.5784
/gene="AGR_C_1406"
CDS
5167.5784
/gene="AGR_C_1406"
/notes="CYTOCHROME C OXIDASE ASSEMBLY PROTEIN CTAG"
/codon_start=1
/transl_table=11
/product="AGR_C_1406p"
/protein_id="AAK86579.1"
/db_xref="GI:15155745"
/translation="MANMEARARAGSQRVSRLVLCVLPFCAMIGMAVAAPLXYS
LFRVGVNGTTRORVQSDVILDKINTVTPDANTSNGNLNMDPFDVKVKEPGETVI
KVTFKAINRFPVAVTGVAFNVTPMEAGAFNKKCECFETETTLQPGTLEMPVFFV
DDPIARETKSVHTLTSYTFYPAKTEREASLPVKTBESGSKL"
5846.6721
/gene="AGR_C_1407"
CDS
5846.6721
/gene="AGR_C_1407"
/notes="(AJ242592) cytochrome c oxidase subunit III"
/codon_start=1
/transl_table=11
/product="AGR_C_1407p"
/protein_id="AAK86590.1"
/db_xref="GI:15155746"
/translation="MADTHOKNDHYHIDPSEWPLLASIGAFITFGGVCVMYRLSGG
SPKLGALANPLFVIGLIVLYVYAWADTIKEANEGSHTRVTVSLHLRYGMIFL
ASVMPFVAFVAFDASLPFHEAIOASRLVGTGOWPKGEIVDPWELPLYNVIL
LLSGTCTWAAHALLHNDKGLLSGLALTVALGVLESTVQVYEHAPDPKNSIYGA
TFMATGTFHGFVGVGVFLVCLFRAIAGGFTPKQHFGEAAWIHFVDVWDLFLF
FAYIMGWGWAPLHG"
6941.7369
/gene="AGR_C_1409"
CDS
6941.7369
/gene="AGR_C_1409"
/notes="(AJ242592) hypothetical protein"
/codon_start=1
/transl_table=11
/product="AGR_C_1409p"
/protein_id="AAK86581.1"
/db_xref="GI:15155747"
/translation="MYSRQPCQREVLMSPEPNSGNFAPVDPIKVLKGCPCPCNGR
LFDGFLTPKACAGLDYGFADAGDPAVFMVLI VGRVLVGMALWQDRPAPPVWAH
VMLWLPVTIVSLVLRKLGIMIALQYRNNASEGRLDRE"
7338.8114
/gene="AGR_C_1411"
CDS
7338.8114
/gene="AGR_C_1411"
/notes="(AF182952) Surfeit 1"
/codon_start=1
/transl_table=11
/product="AGR_C_1411p"
/protein_id="AAK86582.1"
/db_xref="GI:15155748"
/translation="MQAKGLIYNDVSHMKAAQRVWFAAPLVLALAILLGLTWQV
KRLVWKALMADIEERNASPAISIEALVKSQGEIYRKYRLSCTDHTREHFFA
THOQTGYITPLTADGRILFVNRGFPVFEMKEAKRPDQVSGEVTITGLARAPL
VAKPSSLLPNDIAKNIIFYWKLDAAMASSAEVPPERLVKLFVDADNAENPGWPGGV
TLIDLNNHLYQALTWYGLAALVIYAGFAYFNGKAQGE"
8173.9219
/gene="AGR_C_1414"
CDS
8173.9219
/gene="AGR_C_1414"
/notes="(AB006203) LysB"
/codon_start=1
/transl_table=11
/product="AGR_C_1414p"
/protein_id="AAK86583.1"
/db_xref="GI:15155749"
/translation="WNIYVSKPFLTLRLCPGRCAGVDRAIQIVVLLKAYGAPVYV
RHEIVHRYVVEGLKQALFVELHEIPAEHREQVSVFASGVKSPVEDAQANLF
YLDATCPVLVSKVQKQNRHQRHVLVIGHAGHPVIGTMQLPEGVISLVEIVEDA
GVYEPVDRENLFVTTTLSDVDDTAGVIARLQERFPFAIOAPADSDICVATTNRQDAVK

gene
9312.10280
/gene="AGR_C_1416"
CDS
9312.10280
/gene="AGR_C_1416"
/notes="HOMOSERINE KINASE (HK)"
/codon_start=1
/transl_table=11
/product="AGR_C_1416p"
/protein_id="AAK86584.1"
/db_xref="GI:15155750"
/translation="NAVYIDITDEBLRNLFTQYVGSLSYKGAEGVENSFNLLHTT
KPLILITVEKREXNDLPFFGLMQHLAAKSLCPFLPRKDELJUGELSGRPAALI
SLEGMWLEKPEAKREVRKALAAHMLASEGFEIKRPNALSDVQWKLWMDKBERAD
EVEKGLBEIRPEIDYLAHWPDLFAGVIHADFDQNVNFFLGBELSLIDIFYACND
LLAYDVISICLNAWCFEKDQAYNVTKGALLGYSQVPLSFAELEALPLLSRGSALRF
FLRLYDLWLTTPAGALVVKDPLEYLRKLRPHTIANVAEYGLAGE"
10277.10717
/gene="AGR_C_1417"
CDS
10277.10717
/gene="AGR_C_1417"
/notes="Ribbonuclease H (E.C.3.1.26.4) Mutant With His 62
Replaced By Pro (H62p)"
/codon_start=1
/transl_table=11
/product="AGR_C_1417p"
/protein_id="AAK86585.1"
/db_xref="GI:15155751"
/translation="MKHVDIFDGCASNPPGPGWAGVLRVGETEKELSGSEADTTNN
RMELAAI SALNALXSPCEVDLYTDSAYKDGITWIFGWKKKGWKATADNPKRVNVEL
WQALEAAQERHKVTLHWVKGHAGHPENERADELARKMEPFRR"

ORIGIN

Query Match 44.2%; Score 740.4; DB 1; Length 10835;
Best Local Similarity 68.0%; Pred. No. 1.5e-143;
Matches 1109; Conservative 0; Mismatches 511; Indels 12; Gaps 5;
Qy 13 GCATTACGGCCATGACCAACCATGAGCAAGGCTTTCACGCGTGTTCATGTG 72
Db 2025 GACGATCATTCGATCATTCATCCCATGAGCGGGCTTTTCGCCCGCTTCCTGTCC 2084
Qy 73 ACCAACACAGAACATCGTCTGCTATACCTTGTAGCGGTGTGTGTGCTTTCATT 132
Db 2085 ACCAACACAGAGCATCGGACGCTGTACCTGTCTTCGCCATCATGCGGCTATCATC 2144
Qy 133 TCCGTCTGTTCACCTTACATGCGCTTGAGCTGATGGATCCGGGTGTTCAGTACATG 192
Db 2145 GCGCGGGGCTTTCGGTCTGCTATGCGTATGAGCTGCAGGACCGGGCATCCAGATCTTC 2204
Qy 193 TGCCTTCAAGGGCACGCTGATCG---CGATGCTCGCAGACATGTACGGCGAACGG 248
Db 2205 CAGGCTCTGCTTCCATGGTCTACGGCTTCGAGGGGATGCCGCCATCGACGCGGTAA 2264
Qy 249 ACACCTGTGGAACGTATGTTTACCTACCATGATTCCTGTATGATGTTTCTTGTGGGTAT 308
Db 2265 -CATGTTTCAAGTGTTCACGACAGCGACGCGCTGATCATGATCTTCTTTCATGGTCAT 2323
Qy 309 CCGCGCATTTCTCGTGTGTTTGTGTAACCTATCTGATGCGCTGCAATCGGGGCTCCGGA 368
Db 2324 CCGCGCATGATCGCGGTTTTGCCAATCGATGATGCCGATCATGATCGCGCGCTTGA 2383
Qy 369 TATGGCTTCCCGGTATGAACAACTGCTGTTCTGGTGTTCATGCGCGGTACCGCAT 428
Db 2384 CATGGCTTCCCGGCTT-AACTATTTCTTCTGGCTCATCGTTCGCGCTTCATCCT 2442
Qy 429 GGGCGTGGCTTCTGTTTCGACCGGGGTGACGCTGAGCTGGGTTCGGGGTGGTTG 488
Db 2443 GCTGCTCTGCTCTTCTGTCGAGGTCGGGGGTGTTATGCGGTGGGTGGCGCTG 2502
Qy 489 GGTTCGTACCCCGCTGTCGACC---CGGAAGCTGCGCTATTCATGAGCTCGCAT 545

Db 2503 GACCATGATCGCCCTCTGTCGACCAAGTGGATGCCCGGCGCGGTGGATCTTGGAT 2562
Qy 546 TTTCCGGGTTCACTGTGCGGTGCTCTCGATCATGGGCGCGCATCAACATGATCAGAC 605
Db 2563 CTTCTCGTGCAGTGCAGTGCATCGATCCTCGGCGCCATCACTTCAACCCAC 2622
Qy 606 CTTCTTGAATGCGCGCCCGCGCATGACGCTGCACAAAGTCGCTGTTCTTCGTGTC 665
Db 2623 CATCTGAACATGCGCGCGCGCGCATGACGCTGCACAAAGTGCCTGTTGCTGTC 2682
Qy 666 GATCTTTATCAGGCTTGGCTGATCTCTGCGCGCTGCGGTTCTGCTGGTGAATFAC 725
Db 2683 GGTCTGTGTCACCGGATCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2742
Qy 726 CATGCTGTGACCGACCGTAACTTCGGCAGCAGCTTCTTCAATCTCTGCGCGCGGTGA 785
Db 2743 CATGCTTCGACCGACCGTAACTTCGGCAGCAGCTTCTTCTCGCGGAGCGCGGTGA 2802
Qy 786 CCGATTTCTGACCAACATCTCTGTGTTCTTTTGGGACCGGAAATGATCATATCAT 845
Db 2803 TCCGATCTCTTACCAAGATCTGTTCTGTGTTCTTCTGCTTCTGCTTCTGCTTCT 2862
Qy 846 TCTGCGCGGCTTGGCATCATCAGCATCTGCTGTCAGACCTTCTCGAAAGCGGCTT 905
Db 2863 CTTGCGCGGTTTGGCATCATCAGCCATCTGCTGTCACCTTCTTCCAGAGCGGCTT 2922
Qy 906 CGGTTACCTCGCATGCTGATGATGATGTCGATCTGTTCTGGGTTCTGCTGCTG 965
Db 2923 CGGTTATCTCGCATGCTGATGATGATGTCGATCTGTTCTGGGTTCTGCTGCTG 2982
Qy 966 GCGCACCACATGATACCGTGTGATGCTGCTGACCGCAATCTTCTATGCTGCTG 1025
Db 2983 GCGCCACACATGATACCGTGTGATGCTGCTGACCGCAATCTTCTATGCTGCTG 3042
Qy 1026 CACCATGATGATGCGGTCGCGACCGGATTAAGATCTTCTGFGATGCGCACGATG 1085
Db 3043 CACCATGATGATGCGGTCGCGACCGGATTAAGATCTTCTGFGATGCGCACGATG 3102
Qy 1086 GCGCGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1145
Db 3103 GCGCGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3162
Qy 1146 CACGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1205
Db 3163 CACGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3222
Qy 1206 CACACCTATTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1265
Db 3223 CACACCTATTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3282
Qy 1266 GATCTTCCGCGTATCTACTTTTACATGCGGATCTTCCGCGCGCTTCCCGGATG 1325
Db 3283 GATCTTCCGCGTATCTACTTTTACATGCGGATCTTCCGCGCGCTTCCCGGATG 3342
Qy 1326 GCGTCAAAAGTGCATCTTCCGCTTCTTCCGCGCGCTTCCCGGATGATGATGATG 1385
Db 3343 CATCGGCAAGTGCATCTTCCGCTTCTTCCGCGCGCTTCCCGGATGATGATGATG 3402
Qy 1386 GCATCTTCCGCGTATCTACTTTTACATGCGGATCTTCCGCGCGCTTCCCGGATG 1445
Db 3403 GCATCTTCCGCGTATCTACTTTTACATGCGGATCTTCCGCGCGCTTCCCGGATG 3462
Qy 1446 GCTGTGGAACAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1505
Db 3463 CGGCTTGGAACAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3519
Qy 1506 CATCTGATCTTGTGATACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1565
Db 3520 CTTCTCTTCCGCGTCTGCGGAGCTTCTGCGCAAGAGCGCATCGCGGTAACATCTCT 3579
Qy 1566 GCGGAAATTCGCGGATGCTGGAATGAGCGCTGCGGATCAGCGCTTCCGCGCGCTG 1625
Db 3580 GGTGGAAGTTCGCGGATGCTGGAATGAGCGCTTCCGCGCGCTTCCGCGCGCTG 3639

Qy 1626 CGAAACGCTGCC 1637
Db 3640 GGAACGCTGCC 3651

RESULT 10
AB014356

LOCUS
DEFINITION

Brucella suis 1330 chromosome I section 42 of 190 of the complete genome.

ACCESSION
VERSION
KEYWORDS

AE014356.1 GI:23347257

SOURCE
ORGANISM

Brucella suis 1330
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Brucellaceae; Brucella.

REFERENCE
AUTHORS

Paulsen, I.T., Seshadri, R., Nelson, K.E., Eisen, J.A.,
Heidelberg, J.F., Read, T.D., Dodson, R.J., Umayam, L.M.,
Beanan, M.J., Daugherty, S.C., Deboy, R.T., Durkin, A.S., Kolonay, J.F.,
Madupu, R., Nelson, W.C., Ayodeji, B., Kraul, M., Shetty, J., Malek, J.,
Van Aken, S.E., Riedmuller, S., Tettelin, H., Gill, S.R., White, O.,
Salzberg, S.L., Hoover, D.L., Lindler, L.E., Halling, S.M., Boyle, S.M.,
and Fraser, C.M.

The Brucella suis genome reveals fundamental similarities between
animal and plant pathogens and symbionts

Proc. Natl. Acad. Sci. U.S.A. 99 (20), 13148-13153 (2002)

12247741
12271122

2 (bases 1 to 10154)

Paulsen, I., Seshadri, R., Nelson, K.E., Eisen, J.A., Heidelberg, J.F.,
Read, T.D., Dodson, R.J., Umayam, L.A., Brinkac, L.M., Beanan, M.J.,
Daugherty, S.C., Deboy, R.T., Durkin, A.S., Kolonay, J.F., Madupu, R.,
Nelson, W.C., Ayodeji, B., Kraul, M., Shetty, J., Malek, J.A., Van
Aken, S.E., Riedmuller, S., Tettelin, H., Gill, S.R., White, O.,
Salzberg, S.L., Hoover, D.L., Lindler, L., Halling, S.M., Boyle, S.M.,
and Fraser, C.M.

Direct Submission
Submitted (14-AUG-2002) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA

FEATURES
source

1. 10154
/organism="Brucella suis 1330"
/mol_type="genomic DNA"
/strain="1330"
/db_xref="taxon:204722"
/chromosome="I"
135..1034
/locus_tag="BR0463"
135..1034
/locus_tag="BR0463"
/note="identified by similarity to OMNI:NTL02ML6198; match
to protein family HMW PF03755; match to protein family HMW
TIGR00255"
/codon_start=11
/transl_table=11
/product="conserved hypothetical protein TIGR00255"
/protein_id="AA29406.1"
/db_xref="GI:23347258"
/translation="VALQSMGTGFARHAQGMADGEARIVWEVRSVNGKGLDLRLRLP
QGLSEVHPVRSLLARHFRGNFQASLTVERSETQAGFAINQPMALAEVLKGAELQOEK
YGLAPASVDIGILALRGIIIDQSOAREDEBEERALEATIIVAFEGALKAIAEARRQOEK
LFTILSHLDTIERLTADARNPSRATBAIKARLAGOVALLDATREIDETRLYCEAA
FLATKADIQBEIDRLSTHVASARKLLGEGGPGVRKLDLFLSQEFNREANTLCSKSNAS
ITAIGLGLKAAVQDFREQVQNL"

gene
CDS

1043..1705
/gene="gmk"
/locus_tag="BR0464"
1043..1705
/gene="gmk"
/locus_tag="BR0464"

/EC number="2.7.4.8"
/note="identified by similarity to EGAD:13424; match to
protein family HMM PF00625"
/codon_start=1
/transl_table=11
/product="guanylate kinase"
/protein_id="AAN29407.1"
/db_xref="GI:23347259"
/translation="MAISVENVGAVRGLMVVISPSPGAGKSTIARLLLEDPQMKLSL
SISVTRRRSEIDGVHYFTTREFRLNDDELBEAVHGNFYGTIRETAIAL
ADGQMLDIDMOGADLOAKMPADVSIFFILPTMRELOQLNRRAEDTAETVETRL
QNRFEIQKWKYDIVINEDDLRSYAIAKSIINAERLRDRRFGQLDFVVTGLLEEDP
GN"
complement (1732..3144)
/locus_tag="BR0465"
/locus_tag="BR0465"
complement (1732..3144)
/locus_tag="BR0465"
/note="identified by similarity to EGAD:30245; match to
protein family HMM PF01523"
/codon_start=1
/transl_table=11
/product="tldD protein, putative"
/protein_id="AAN29408.1"
/db_xref="GI:23347260"
/translation="MKSILDSPTGCRDHIERLVAESLKGSDDGELFLEYREAEALVFD
NRLKTGFGNQCGFLRSVAGEAIGYAHAGDLSLAALKRASDAVSAILKYGCTYSP
APAGNTRHLYTDENFGIPGDFDAKVQLLOQIDAYLRADKTKVRSVSLACSQQVEI
LRADGHEVDIRPMARLSVSVAGSDQSGSYTLGGRKGFGFETTESWOHADEA
LRAQVNLAEAIAPAGTDFIVLANGWPGVMLHEAVGHLEGDENRKTTSAFAGLIGQQ
VAATGVVDDQTIATERRSLTIDDEGTPTNRTVLIEDGKLVVYVQNRQNALMGMEA
TCNGRRSVAHAPMPRMUNTYMLSGDKTPEETIASIRKGIYAVSFGGQVDITSGKFA
FQCTEAYMIEDQKVAPIKGMALINGDPDMKEIISMGNDMKLDTGSCNCGKGQWVP
VVGQPHLRMDQITVGGTAV"
complement (3286..3882)
/locus_tag="BR0466"
complement (3286..3882)
/locus_tag="BR0466"
/note="identified by similarity to GP:15073834"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAN29409.1"
/db_xref="GI:23347261"
/translation="MTQRLAALAAAGLGLSLAAATLPSAONNTPGTQQLPAPQOQTPA
EQPTTSAPQSGALKSOHAGWSILCDTPAGAKTEOCALIONVAAKPELGLSVV
LKTADNKAILRLVLAELGVLNGLNVDGKIDGRXVRCFEDGCGYAEVLEDDLL
KTLRNGKATFIVFQTPREGVGIPVDLKGFGEGFDALP"
4346..5200
/gene="coxB"
/locus_tag="BR0467"
4346..5200
/gene="coxB"
/locus_tag="BR0467"
/EC number="1.9.3.1"
/note="identified by similarity to EGAD:8295; match to
protein family HMM PF00116; match to protein family HMM
PF02790"
/codon_start=1
/transl_table=11
/product="cytochrome c oxidase, subunit II"
/protein_id="AAN29410.1"
/db_xref="GI:23347262"
/translation="MGLAVLMTTGAALADQPPWRFQDAATGAEQIHFERYT
LWFTPTLVLGLVWVLRFRASVNPSPKTSHTNTAEIVMTVGPVILLFLVAPS
FOLLTAQYSPEDPLTVKATGYQWYGVQVNDPVSDVIRKDDDAAGKEDRAR
YPLRLVDNREVPVGETVELLVTASDVTHSWTIPAFGVKMDAVPGRINETWFKADKE
GLYYGCELCGDKDGHFMPIAVRVSKDQVQNWLSAAATDLPGANKALLQAEGBSSAN
DKVAAGL"
5338..6996
/gene="coxA"
/locus_tag="BR0468"
5338..6996
/gene="coxA"

/locus_tag="BR0468"
/EC number="1.9.3.1"
/note="identified by similarity to EGAD:5976; match to
protein family HMM PF00115"
/codon_start=1
/transl_table=11
/product="cytochrome c oxidase, subunit I"
/protein_id="AAN29411.1"
/db_xref="GI:23347263"
/translation="MAGTAHAEHGAHDDHKPGVWVWVSTNKHDTGLTLYLPAITAG
IIGGALSIAMRAELOEPQIGIQIFHLQAMVYEGDADAGKRMFNVFTSAHALVMIF
FYMPALIGFANGWVPLMGADPMAPFRMNNISFWLLPPALTLIIILFMPGPAGGW
GGGNTLTPPRSTSGPGPVDFAIHAHSGASSILGAINFTITLNNRPAQMTHLH
KPFLEAVNLGTAFILLLSIPVLAGITMLTDNRNFTTFFAPDGGDDPILYOHLETVG
FGPEVYILLPGFGIVSHVSTFSKPIEGYLGMAVAVAGVGVGVVVAHMYTVG
LSLDTQRYFVFTAMVAVPTGKIFSWIATIMWGGSSISPRAPMLMAGFIFLPIGVT
GVLANAGLDRALHDITYVVAHVYLSIGAVFAIFAGWYIFPPKWSGTMINEFTAKL
HFVTVFVGNLVFFPQHFLGALPPRYIDYDAYAGWNNISSYYSISFAVLIELY
NVFEAKKREAGANPGDGAATLEWQLSSPPPFHQQEQLPRIK"
7224..8171
/gene="ctaB"
/locus_tag="BR0469"
7224..8171
/gene="ctaB"
/locus_tag="BR0469"
/EC number="2.5.1.-"
/note="identified by similarity to EGAD:5818; match to
protein family HMM PF01040; match to protein family HMM
TIGR01473"
/codon_start=1
/transl_table=11
/product="proteoheme IX farnesyltransferase"
/protein_id="AAN29412.1"
/db_xref="GI:23347264"
/translation="MSLVEKNTASEDAFALSEATARDYLVLKPRVMSLVVFTGLVGL
VLAPGHNPVIAVISILICIAVGAGSALNMWYDADIDAVKSTRKRPPIAGIAPNQ
VLAPGLTSAFVSNTIGLMVNWLAALALFTIFFYAVIYTMWLKRSTPQNVIGGAG
AFPPMIGMAATGEITWDSLVLFMIIFLTPPHFWALSLETTINDYEARIIPMPFNKVG
ELSTRQALFYALVPLWGVLPWVGFAGVYGVVSTLLGLAFVYVYAWRLWAADSPQ
MLAARKLFRFSLLYLAGIEFAVLLFEALTKEKLLAAFGVF"
8171..8356
/locus_tag="BR0470"
8171..8356
/locus_tag="BR0470"
/note="identified by similarity to OMN1:NTL02ML5908"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAN29413.1"
/db_xref="GI:23347265"
/translation="MDKTAHKPVHEKLELVVPTPEQKKAORNSWGLAIALALFVLIV
YIGTLAKYAGVFMRP1"
8376..8981
/gene="ctaC"
/locus_tag="BR0471"
8376..8981
/gene="ctaC"
/locus_tag="BR0471"
/note="identified by similarity to EGAD:33787; match to
protein family HMM PF04442"
/codon_start=1
/transl_table=11

Query Match 44.1%; Score 738.2; DB 1; Length 10154;
Best Local Similarity 66.9%; Pred. No. 4.3e-143;
Matches 1097; Conservative 0; Mismatches 533; Indels 9; Gaps 3;
QY 5 CAGAGCGCGCATTCACGCGCCATGACCACCATGAGAACCAAGGCTTCTTTCACGCGCTCGT 64
DB 5348 CAGCAGCTCACGAGCATGTGTCGCCACGAGCACCACCAAGCCGATCGCTGGGTCTGTTGGG 5407
QY 65 TCATGTGCAACCAACCAAGACATCGGTCTGCTATACCTTGTACCGGCTGCTGTTGTTG 124

5408	TAT	ACTCGACCAATCATATAAGACATCGGTACCCCTGTACTCTGATTTTTCGAATCATCGCG	5467
125	GT	TTCAITTCGGTCTGTTCACTACCGCTCTACATGCGCCTTGAGCTGATGGATCCGGGTGTC	184
5468	GC	ATCATCGCGCGCGCTTTCCATCGGATGCGTCCGAACTCCAGAGCGGGCATCC	5527
185	AG	---TACATGTGCTTGAAGCGCGCATGTCTGATTCGGGATGCTCTCGAGACATGTACGG	241
5528	AG	ATTTTCCATGTGTTGGCCAGATGGTTTACCGTGTTTGAAGCGGATCGCGCCCTTGATG	5587
242	CG	AAACGGACACCTGTGAAACGTATGTTACTTACCATGGTATTCTCATGATGATGTTCTTTG	301
5588	CC	GCAGACAGCATGTTCAAAGTGTACCTCCGCGCAGCGCTCGTCAATGATCTTCTTCA	5647
302	TG	GTATCCCGCGAATGTTTCGTGGTTTTGGTAATCTATCTGATGCGCTGCAAATCGGCG	361
5648	TG	GTATCGCTCGCTCATCGCGGTTCGCCAACTGGATGGTGGCTCGCCTCATGATCGCG	5707
362	CT	CGGATATGGCTTCCCGGTATGAACAACCTGTCGTTCTGGCTGTTCAATGCGGTA	421
5708	CG	CGGACATGGCGTTCCTCAGCAGTAAACAAATATTCCTTCTGCTCTGCCGCCCGCGC	5767
422	CC	CGATGGCGGTGCTTCGCTGTTCACACCGCGCGGTGACGGTCACTGGTTCGCGCG	481
5768	TG	AGCTGCTGCTGATCTCGCTGTTATGCCCGTCCGCAGGTGGCTGGCGCCCGCGCG	5827
482	TT	GGTTCGGGTTCGTACCGCGCGCTGTGCA---CCCGGAAGCTGGCTATTCATGTGACC	538
5828	CG	CTCGACGCTCTATCCGCGGTTCGACACTCCCGCCAGCCGCGCCCGCGGTGGATT	5887
539	TC	CGATTTTCGGGTTCACTTGTGCGGTGCTCCTCGATCATGGCGCGCATCAACATGA	598
5888	TC	GCATTCGGGCATCCACATTTCCGCTGCCCTCGTCACTCTCGCGCGCATCACTTCA	5947
599	TC	AGACTTTTGAAATCGCGGCCCGCGCATGACCTGTCACAAAGTCCGCTTGTCT	658
5948	TT	ACGACGATCCTGAAATCGCGCGCCGGCATGACCTGTCACAAAGTCCGCTTGTCTG	6007
659	CG	TGTCGATCTTTATCAGCGCTTGGCTGATCTGCTGGCGTCCGCTTCTGGCTGTG	718
6008	CT	TGGCCGTTCTGGGTACGCGCTTCTGCTGCTCTGTCTCGGTTCTCGCAGGTG	6067
719	CA	ATCAACATGCTGCTGACCGACGTAATTCGCGACAGACTTCTTCAATCTCTGCTGGCG	778
6068	GT	ATCAACATGCTGCTTACCGACCGCAATTCGCGACAGACTTCTTCCGCGCTGATGGCG	6127
779	GC	GTACCGATTCGTACCAACACATCTGTGGTTCCTTGGGACCCGGAGTGTACA	838
6128	CG	CGGACCCGATCTCTACAGCACTGTTCGTTTCTGGCCATCCGGAAGTCTATA	6187
839	TC	ATCATCTTGCCTCGGCTTTGGCATCATAGCCATGCTGTTGACCTTCTCGAAAGC	898
6188	TC	CTGATCTCTCGCGGCTTCGGCATCGTCAAGCCATGTGTCTCCACTTCTCGCGCAAC	6247
899	CG	GTCTTCGGTTACTCTCGCATGGTCTATGCAATGGTGGCAATCGGTGTTCTGGGCTTTG	958
6248	CG	ATCTTCGGTTATTCGGCATGGCTTAGCCCATGGTGCAATCGGTGTTCTGGGCTTTG	6307
959	TC	GTCTGGGCGCACCATGTACACCGTTGGTATGTCGCTGACCCAGCAATCTCACTTCA	1018
6308	TC	GTCTGGGCCACCATGTATACGCTGGCCCTTTCGCTCGATACACAGGCTACTTTCG	6367
1019	TG	CTGGCCACCATGTGTATCGCGGTCCGACCGGCATTAAAGATCTTCTCGTGATCGCA	1078
6368	TC	TTCGCCACATGGGTCAATCGCGGTCCGACAGGCATCAAGATCTTCTCGTGATCGCA	6427
1079	CG	ATGTGGGCGGCTCGGTGTAGTTCAAAATCGCGCATGCTCTGGGCTTTTGGCTTTATGT	1138
6428	CG	ATGTGGGTGACTCGATTTCTTTCGGCGGCCCAATGCTCTGGGCTATCGGTTCACT	6487
1139	TC	GTCTCACCGTGGGTGGTGTGACCGGTATCGTGTGCGCCCAAGGGGTCTGGACCGTG	1198
6488	TC	GTCTCAACATTCGCGGTGTGACGGGTGTGAGCTTCGAAATGCGGTCTCGACCGTG	6547

QY	1199	CATATACGACCACTATTACGTGGTGGCGCACTTCATATTATGATATGCTCGCTGGGTGCGA	1254
Db	6548	CAITGACAGATACCTATTATATGTGGTGGCGACATTCACACTATGTCTGTGCTGTGGCGCTG	6607
QY	1259	TCATTGCGATCTTCGCGGATATCTACTTTTACATGCGGAAGTTCTCGGGCGCGCTTTCC	1318
Db	6608	TCITCGCGATCTTTCGCGGCTGGTACTACTGTTCCGAGATAGAGCGGTACATGTATA	6667
QY	1319	CGGAATGGGCTGCAAGCTGCACCTCTGAGACCTTCTTCATCGTGGCGAAAGTCACTGTTCT	1378
Db	6668	ACGAGTTTCACGCGAAAGCTGCATTTCTTGGGTGCGTTCAGTTCTGTCGCTGTGAACCTGCTCTCT	6727
QY	1379	TCCCGCAGCACTTCTCTGGGACGTCAGGGTATGCGCGCGCTTACATCGACTATCCCGAAG	1438
Db	6728	TCCCGCAGCACTTCTCTGGGCTCGGGTCTCGCGGCAATGCGCGCTTACATCGACTATCCCGATG	6787
QY	1439	CTTTTCGGCTGTGGAACAAAGTCTGCTTCATGTTGGTTCCTGGGCTTCGCTCGCTGTTCC	1498
Db	6788	CTTATGCGCGCTGGAACATGATTTTCGTCATATGGCTCCTACATTTACGGCTTTGCCGTGC	6847
QY	1499	TGTTCTTCATCGTGATCTTTGTCTATACGCTGCTGCTGCGCGCGCGAGACCGTCCGA	1558
Db	6848	TGATCTTCTCTATAAC--GTCTTCAGGCGCTTCGGAAGAACGTTGAACGGGTGCAA	6904
QY	1559	ACCGTGGGCGCAATTGCGCGATACGCTGGAATGAGCGTGCATCACCGCTTCCGCGCC	1618
Db	6905	ATCCGTGGGCGATGCGCGACCACTGGAATGGCAGCTTTCTTCGCGCGCGCTTCC	6964
QY	1619	ACAGCTTCGAAACGCTGCC	1637
Db	6965	ACCAGTGGGAACAGCTCCC	6983

RESULT 11
AE009583/c

LOCUS
Brucella melitensis 16M chromosome I, section 140 of 195 of the complete sequence.

ACCESSION
AE009583

VERSION
AE009583.1

KEYWORDS
GI:17983463

SOURCE
Brucella melitensis 16M

ORGANISM
Brucella melitensis 16M

REFERENCE
AUTHORS
DelVecchio,V.G., Kapratl,V., Redkar,R.J., Patra,G., Mujar,C., Los,T., Ivanova,N., Anderson,I., Bhattacharyya,A., Lykdis,A., Reznik,G., Jablonski,L., Larsen,N., D'Souza,M., Bernal,A., Mazur,M., Goltsman,E., Selkov,E., Elzer,P.H., Hagius,S., O'Callaghan,D., Letesson,J.-J., Haselkorn,R., Kyrpides,N. and Overbeek,R.

TITLE
The genome sequence of the facultative intracellular pathogen *Brucella melitensis*

JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 99 (1), 443-448 (2002)

PUBMED
11756688

REFERENCE
AUTHORS
DelVecchio,V.G., Redkar,R.J., Patra,G. and Mujar,C.

TITLE
Direct Submission

JOURNAL
Submitted (13-NOV-2001) Institute of Molecular Biology and Medicine, University of Scranton, Scranton, PA 18510, USA

REFERENCE
AUTHORS
Elzer,P.H. and Hagius,S.

TITLE
Direct Submission

JOURNAL
Submitted (13-NOV-2001) Department of Veterinary Science, LSU Ag Center, 111 Dalrymple Building, Baton Rouge, LA 70803, USA

REFERENCE
AUTHORS
Kapratl,V., Los,T., Ivanova,N., Anderson,I., Bhattacharyya,A., Lykdis,A., Reznik,G., Jablonski,L., Larsen,N., D'Souza,M., Bernal,A., Mazur,M., Goltsman,E., Selkov,E., Haselkorn,R., Kyrpides,N. and Overbeek,R.

TITLE
Direct Submission

JOURNAL Submitted (13-NOV-2001) Integrated Genomics, Inc., 2201 W. Campbell
Park Drive, IL 60612, USA
REFERENCE 5 (bases 1 to 11404)
AUTHORS Letesson,J.-J.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-2001) Unite de Recherche en Biologie Moleculaire,
Laboratoire d'Immunologie et de Microbiologie, Universite of Namur,
61 rue de Bruxelles, Namur 5000, Belgium
REFERENCE 6 (bases 1 to 11404)
AUTHORS O'Callaghan,D.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-2001) Faculte de Medecine, INSERM U431, Avenue
Kennedy, Nimes 30900, France
FEATURES
source
1..11404
/organism="Brucella melitensis 16M"
/mol_type="genomic DNA"
/strain="16M"
/db_xref="taxon:224914"
/chromosome="1"
complement(85..834)
/gene="BMEI1460"
complement(85..834)
/gene="BMEI1460"
/codon_start=1
/transl_table=11
/product="SURPELT LOCUS PROTEIN 1"
/protein_id="AAL52641.1"
/db_xref="GI:17983464"
/translation="MTLQTSRRFPWGLVLSGIALVLLSGTWQVRLMKKALLA
STQRHPPEPLAEEMKIQEGTVEYRPTVSGTFHQGERFLATYKGEAGFYV
TEPLEDRGRVLRGVEYRKQDPATRPAGELAGPVTGLARDPLSKVPSFLVPD
NDIAKNI FYKMDMAAESAGLPDLGRVVPFVDADSPNPGGLPIGVTTIDFPNNH
LQAYTWGLALALGVGTWLMRYKSGRG"
complement(831..1289)
/gene="BMEI1461"
complement(831..1289)
/gene="BMEI1461"
/codon_start=1
/transl_table=11
/product="ZINC-FINGER PROTEIN"
/protein_id="AAL52642.1"
/db_xref="GI:17983465"
/translation="MAAFCSQVAFQSRNCRCFRGRKNAMPEHGHYPQVDPIRSGIAG
RCRCGEGKLFNGFLVTPACGLDYGPADSGPAAFLIIGFVVGWALMVEV
NTGPLVWHFILWPLAILSLVAMRWIKILINQHRKAEQCRGLDSNK"
complement(1328..2272)
/gene="BMEI1462"
complement(1328..2272)
/gene="BMEI1462"
/EC_number="1.9.3.1"
/codon_start=1
/transl_table=11
/product="CYTOCHROME C OXIDASE POLYPEPTIDE III"
/protein_id="AAL52643.1"
/db_xref="GI:17983466"
/translation="WRFTLQAGSAREEPLRGNTGRMDADVHOKNDHYHIIAPSPFPL
SSVGFVLAIGIAMWRYINAGELPFFGLNVVTPWIFGLAIYVLCFMGWSDTIKE
GHEGHTRVSLHLRGIMTFIASEVMFFVAFWAFDASLPFNEHVQVRAEFTFGV
WPKGTGLDPLHLPLNTYVILLSSGTLTWAHALLHNDKRGMITGLAITVALGVLF
SCVQAVEYTHAPFKDSITYGATFFMATGFHGFHVLIGTIFLLVCLFRAIGDFTPKQ
HGFGEAAAYWHFVDVWVLFLEAFIYVWGSGAPMEGG"
complement(2415..3128)
/gene="BMEI1463"
complement(2415..3128)
/gene="BMEI1463"
/codon_start=1
/transl_table=11
/product="CYTOCHROME C OXIDASE ASSEMBLY PROTEIN CTAG"
/protein_id="AAL52644.1"
/db_xref="GI:17983467"
/translation="MPSRRSRFCFWFTSARWRPAPACSGRSDNRIBERMTDQGENE
KKQRNSRNATIAVACLSPFFVCMIGAASVAPVLYRIFCQVTGYGTTQRVEQSDTILDK

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

TIKVRFDANIANGLPWDFKPMOREVTVRIGETTMIKYEAHNLFGBETGRASFNVAPG
RAGAVNKVCEFCFTDNTLKPGEDLELPWFVRFVDFVNDPLDKVKITLITLSTYTFPI
DAPKPVNAKAVGSTRNGG"
complement(3225..4172)
/gene="BMEI1464"
complement(3225..4172)
/gene="BMEI1464"
/EC_number="2.5.1.1"
/codon_start=1
/transl_table=11
/product="PROTOHEME IX FARNESYLTRANSFERASE"
/protein_id="AAL52645.1"
/db_xref="GI:17983468"
/translation="MSLVKNTASDEDAFALSEATARDYLVLKPRVMSLVVFTGLVQL
VLAPGHMNPVLAVISILCIAVGAGSAGALNMWYDADIDAVMKRTRKRPIGAGIIAPNQ
VLARGLTISAFSVMTLGLMVNLAALIAFTIFFVAVIYTMWLKSTPQNIIVIGGAAG
APPMKTGMAATGEITWDSLVLFPMIFLWPPHFVALSFTTNDYEAARIIPWNPVKG
ELSTRQALFYAVLAPVGVLPWVGPMGFYGVUSTLGLAFVYVAVWLWAADSQPG
MLAARKLFRSLYLAGIFAVLLFEALTFKLLAAFGVF"
complement(4400..6058)
/gene="BMEI1465"
complement(4400..6058)
/gene="BMEI1465"
/EC_number="1.9.3.1"
/codon_start=1
/transl_table=11
/product="CYTOCHROME C OXIDASE POLYPEPTIDE I"
/protein_id="AAL52646.1"
/db_xref="GI:17983469"
/translation="MAGTAAGHGHADHDKPHGVWRVYSTNKHDKITGLYLIPAILAG
TIGGALSIAAMRAELQEPQIQLHGLAQVYVGVGDAALDAGKHMNVFTSAHALVMIF
FWMFALIGGFANWYPLMIGAPDMAFPMNNISFWLLPALTLILLISLFMPGPGAGW
GGGGTLLVPPSTSGQGPAPVDFAILAHISGASSILGALNFIITILLNKRAGMTLH
KMFLEAWLVTAFLLLSLPLVLAGITLWLLDNFTGTTFFAPDGGDPILOYHLEWF
FHPVEVITIQPGFIVSVHVSFTRKPIFGYLGMAVAVAGVGVGVVVAHHMYTVG
LSLDTQRYFVFATVIAVPTGIKIFSWIATMGGISFRAPMLWAIQFTLFTIGVY
GVOLANGLDRALHDTYVVAHPHVLISGAVFAI PAGWYMPKMSGYMNEFTAKL
HFWTFVGNLVFFPQHFGLGMPREYIDYDPAVAGMNMISYKSGVISGFAVLIFY
NVFEAFKKEAGANFGWGAGTLEWQLSSPPPHQWEQLPRIK"
complement(6196..7158)
/gene="BMEI1466"
complement(6196..7158)
/gene="BMEI1466"
/EC_number="1.9.3.1"
/codon_start=1
/transl_table=11
/product="CYTOCHROME C OXIDASE POLYPEPTIDE II"
/protein_id="AAL52647.1"
/db_xref="GI:17983470"
/translation="MNADGKAVLKTGVQPERAKRRLVVDKIYAATRGAGGALAVL
WTTGAALADOPRWEKRFQDAATGIAEQIHWFRYTLWFTIITLLVLGLVWVLRFF
RASVNPSPKTSHTALEVITVGVPIILLFLAVSPOLLTAQYSPEDPTLTVKATGY
QWNGYBYQVDNPFSDVILKDDDDRAAKGEBRARPRLTVDNWWVPVGTVRLL
VTASDVIHSWTIFAFGVNDVAFGRINETWFKADKGLYTGQCSSELCCKGHGMPFIIV
RVVSKDQYQNWLSAATDLFGANKALLQAESEGSANDKVAAGL"
7514..8110
/gene="BMEI1467"
7514..8110
/gene="BMEI1467"
/codon_start=1
/transl_table=11
/product="Hypothetical Protein"
/protein_id="AAL52648.1"
/db_xref="GI:17983471"
/translation="MIORAALAAAGLSLAATPLPSAQNNTPTGQLPAPQPTPA
EPQPTTSATPQSGALSQSGANSILCDTPAGAKTEOCALIQNVAAKRPGLSVVY
LKTADNKARILIRVLAPLPLNGLGNVDKDLGRAYFVRCFEDGCGYAEVILBDDLL
KTLNKGAAATFVFTPEEGVGIPVDLKGFGEGFDALP"
8251..9663
/gene="BMEI1468"
8251..9663
/gene="BMEI1468"
/codon_start=1

```

/transl_table=11
/product="TLDD PROTEIN"
/protein_id="AAL52649.1"
/db_xref="GI:17983472"
/translation="MKSLIDSPDTRDRIHRIVAESLKSDDGELFLEVREBALVFD
NGLKTSFQDQFGKRSVAGEAIGYAHAGDLSLAKRKSDAVSATLKGKGYTSP
APAGTNHLITDENIPGAPFDKQQLQOIDAYLRADTKVQSVSLACSQQOYVEI
LRADGHVRDIRPMARLSVAGSDRQBSGYTLGRRKGFGEFTTTSWOHAADEA
LRQALVLEAI PAPAGTDFIVLANGPVGMLHEAVGLEGDFNRKRTSAFAGLLQOQ
VAARGVTVDGIIAERRSGLTDDGTPNRTVLIEDGKLFNMODRONRNLGMGEA
TNGRRSYAHAPRMTNTYMLSGDKTPEELIASIRKGIYAVSFGSGVDLITSGKEV
FGCTEAYMIEDGKVGADIKAMLLGNPGDPAKRLISMLGNMDKLDITSGNCGGQWVP
VGVQPHLRDQITVGGTAV"
complement(9690..10457)
/gene="BMEI1469"
CDS
complement(9690..10457)
/gene="BMEI1469"

Query Match      44.08; Score 736.6; DB 1; Length 11404;
Best Local Similarity 66.98; Pred. No. 9.1e-143;
Matches 1096; Conservative 0; Mismatches 534; Indels 9; Gaps 3;

Qy 5 CAGACGGCGGCATTCACGGCGCATGACCACCATGAGAAGCAAGGCTTCCTCACGCGCTGGT 64
Db 6048 CAGCAGCTCAGAGCATGGTGCCACGACGACACAGCGCATGGCTGGGTTGG 5989

Qy 65 TCATGTCGACCAACCAAGACATCGGTCTGCTGTATACCTGTAGCGGTGGTGTGGT 124
Db 5988 TATACTCGACCAATCATAAAGACATCGGTACCTGTACCTGATTTTGAATCATCGCG 5929

Qy 125 GTTTCATTTCCGTCCTGTTCAACCGCTTACATGCGCGCTTGAGCTGATGGATCGGGGTTC 184
Db 5928 GCACCATCGCGGGCGCTTTCCATCGCATGGTGGCGGACTCCAGAGCGCGGATCC 5869

Qy 185 AG---TACATGTGCTTGAAGGCGCAGCTGTGATCGCGGATGCTCGCAGACATGACG 241
Db 5968 AGATTTTCCATGTGTGGCGCGAGATGTTTACGGTGTGAAAGCGATGCGGCCCTTGATG 5809

Qy 242 CGAACGACACCTGTGGAAGCTCATGGTTACCTACCATGGTATCTGATGATGTTCTTTG 301
Db 5808 CCGGCAAGCACATGTTCAAGCTGTTCACTCCGCGCAGCGCTGTCATGATCTTCTTCA 5749

Qy 302 TGGGTATCCCGCATTTGTCGGTGGTGGTAACTATCTGATGCGCGCTGCATAATCGGCG 361
Db 5748 TGGTCATGCTGGCTGATCGCGGGTTCGCCAATGGATGGTGGCGCTCATGATCGGCG 5689

Qy 362 CTCGGATATGGCTTCCCGCGTATGACAACTGCTGCTTCTGGCTGTTTATGCGGTA 421
Db 5688 CGCGGACATGGGCTTCCACGCGATGAACAAATTTCTTCTGGGCTGCGGCCCGGCG 5629

Qy 422 CCGCGATGGCGGTGGCTTGGCTTTCGCACCGGGCGGTGACGCTCAGCTGGGTTCGGCG 481
Db 5628 TGACGTGCTGTGATCTCGCTGTTTCATGCCGCTCGGCGAGTGGTGGCGCCCGCG 5569

Qy 482 TTGGTTGGTTCGTACCGCGCTGCGA---CCGCGAAGCTGGCTATTCGATGGAC 538
Db 5568 GCGGCTGGAGCTCTATTCGCGCGTTCGACCTTCGCGCGAGCCCGCGCGCGGTGGATT 5509

Qy 539 TCGCGATTTTCGGGTTCATTTGTCGGGTGCCTCTCTCGATCATGGCGCGATCAACATGA 598
Db 5508 TCGCCATTCGCGCATCCCATTTTCGGTGGCTCTGTCGATCTCGGCGGATCACTTCA 5449

Qy 599 TACGACCTTTTGAACATGCGGCGCCCGCGATGACGCTGCACAAAGTGGCTGTGTCT 658
Db 5448 TTACGACGATCTCTGAACATCGCGCGCGCGATGACGCTGCACAAAGTGGCTGTCTCG 5389

Qy 659 CGTGGTGCATCTTATACGCGCTTGGCTGATCTGCTGCGCTGCGGCTTCTGGCTGGT 718
Db 5388 CTTGGGCGCTTCTGTTAGGCGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5329

Qy 719 CAATACCATGCTGTGACCGACCGTAATTCGCGACGACCTTCTCAATCCTGCTGGCG 778
Db 5328 GTATACCATGCTGTACCGACCGTAATTCGCGACGACCTTCTTCGCGCTGTATGCG 5269

```

```

Qy 779 GGGGTGACCCGATTCGTACCAACACATCTCTGTGGTCTTTTGGCACCCGGAAGTGACA 838
Db 5268 GGGGCGACCCGATCTCTTACGACACCTGTCTGTGTTCTTCGGCATCCGGAAGTCTATA 5209

Qy 839 TCATCATTTCTCCCGGCTTTTGGCATCATGACCCATGTCGTGTCGACCTTCTCGAAAAAGC 898
Db 5208 TCTCATCCAGCGGCTTCGGCATCGTCAGCCATGTCGTCTCCACCTTCTCGCGCAAGC 5149

Qy 899 CGGCTTCGGTTCATCTCGCGATGCTCTATCAATGTTGGCAATCGGTGTTCTGGGCTTTG 958
Db 5148 CCAATCTTCGGTTCATCTCGCGATGCTCTACGCCATGTTGGCAATCGGTGTTCTGGGCTTTG 5089

Qy 959 TCGTCTGGGCGCACCATGATACACCGCTTGTATGTCGCTGACCCAGCAATCTCTACTTCA 1018
Db 5088 TCGTCTGGGCGCACCATGATACACCGCTTGTATGTCGCTGATACACAGCGCTACTTCG 5029

Qy 1019 TCGTGGCCACCATGGTATCGCGTGGCGACCGCATTAAGATCTTCTCTGGATCGCCA 1078
Db 5028 TCTTGGCCACCATGGTATCGCGTGGCGACCGCATTAAGATCTTCTCTGGATCGCCA 4969

Qy 1079 CGATGTGGGCGGCTCGGTTGAGTTCAAAATCGCGATGCTCTGGGCTTTGGCTTTATGT 1138
Db 4968 CGATGTGGGCTGGCTCGATTTCTTCCGCGCGCATGCTCTGGGCTATCGGCTTCACT 4909

Qy 1139 TCCTGTTACCGTGGTGGTGTGACCGGTATCGGTGTCGCGCCCAAGCGGTCTGACCGTG 1198
Db 4908 TCCTGTTACCATCGCGGTGTGACCGGTGTGCAAGTTGCAAAATGCGCGTCTCGACCGTG 4849

Qy 1199 CATATCAGCACACTATTAGCTGTGGCGACTTCCATTTATGATGTCGCTGGTGGTGG 1258
Db 4848 CATTCGACGATACCTATTATGTTGGCGCACTTCCACTATGCTGCTGCTTGGCGCTG 4789

Qy 1259 TCTTTCGATCTTCGCGGTATCTACTTTTACATGCCGAAAGTTCTCGGGCGCGCTTTC 1318
Db 4788 TCTTTCGATCTTTCGCGGTGGTACTACTTGGTTCGCGAAGATGAGCGGTACTATGATA 4729

Qy 1319 CGGATGGGTGCAAGCTGCACTTCTGGACCTTCTTCATCGGTGCGAAGTCACTTCT 1378
Db 4728 ACGAGTTACCGCAAGCTGCACTTCTGGGTACGTTTCGCGGTGAACTGGTCTTCT 4669

Qy 1379 TCCCGCAGCACTTCTCGCGACGTCAGGATGTCGCGCGCTTACATCGACTATCCGGAAG 1438
Db 4668 TCCCGCAGCACTTCTCGCGCTCGCGGATGTCGCGCGCTTACATCGACTATCCGGAAG 4609

Qy 1439 CTTTCGCGCTGTGGAACAAAGTCTGCTCTATGTTGGTGTCTTGGGCTTGGCTTGGTTC 1498
Db 4608 CTTATGCGGCTGGAACATGATTTGCTCATATGGCTCTTACATTTTTCAGGCTTTTGGCTGC 4549

Qy 1499 TGTTCCTTCATCGTATCTTGTGTATACGCTGGTGTGTCGCGCGCGAGACCGCTCGGA 1558
Db 4548 TGAATCTTCTCTATAAC---GTCTTCGAGCGGTTCGCGAAGAGCGTGAAGCGGTGCAA 4492

Qy 1559 ACCGCTGGGCGAATTCGCGATACGCTGGAATGAGCTGCGCATCACCGCTTCCGCGCC 1618
Db 4491 ATCCGTGGGCGATGCGCGACCAACGCTGGAATGCGAGCTTCTTCGCGCGCGCGCTTCC 4432

Qy 1619 ACAGTTTCGAAACGCTGCC 1637
Db 4431 ACCAGTGGGAACAGCTCCC 4413

```

```

RESULT 12
AP003011 346510 bp DNA linear BCT 15-MAY-2001
LOCUS Mesorhizobium loti DNA, complete genome, section 18/21.
DEFINITION AP003011 BA000012
ACCESSION AP003011.2 GI:114026664
VERSION
KEYWORDS Mesorhizobium loti
SOURCE Mesorhizobium loti
ORGANISM Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Phyllobacteriaceae; Mesorhizobium.

```


gene	/protein_id="BAB53268.1"		358	GGCGCTCCGGATATGGCTCTCCCGGATGAACACCTGCTGTTCTGGCTGTTTCAATGCC	417
	/db_xref="GI:14026672"		337522	GGCGGCCCGACATGGCGTTCGGCGCATGAACAAATCTCTCTTCTGGCTGCTCCGCGCG	337581
CDS	/translation="MATLSIKDVRKSYGVVLHGVLDIELDGGFLVLLPGSCGKST LLNLIAGLDDTSQGDILIGRSVVDLAPKDRNIAMVPSVALYPTVSVERNIGFLGEM RGVDADKRAVDAAQLQLSHLLDRPANSQGORVAMGRAIVDFPELFLFDEP LSNDAKLVEKTEIKRLHERLGTIVYVTHDOIEMTLATKAVMKGGHIOQFADP RTIYERPAVFASFTGSPAMTFLPGLVSDATGVRFAAGSVVALQAGKRTVDSRP VILGIRPEELKVGDRDASVTGVIDVVEPTGPTMTVTQVGDQSIILRLPPAGRRN EPIFLAVDPASVNLFDPSBHRIDI"		418	GGTACCGCGATGGGCGTGGCTTGGCTGTTTCGACACCGCGCGGTGACGGTTCAGCTCGGTTGG	477
	/db_xref="GI:14026673"		337582	GCTTTCATCTGCTGCTCACTCACTGATGTTCTGTCGGAGTGGCGCGCTTCGGCGTC	337641
gene	/gene="msr7097"		478	GGCGTTGGTGGGTCTGTACCGCGCGCTTCGACCCCGGAGTGGCGGCTTCGATG	534
	/db_xref="BAB53269.1"		337642	GGCGCGCGCTGGACGCTCTATCCGCGCTGTCGACCTCGGGTACGCCGACCGCCCATG	337701
CDS	/note="unknown protein"		535	GACCTCGCGATTTTCGCGGTTCACCTTGTTCGGGTGCTCTCTCGATCATGGGCGCGCATCAAC	594
	/codon_start=1		337702	GATCTGGCGATCTCTGCGATCCACATTCGCGGTGCTCTCGATCTCTCGGCGCATCAAC	337761
gene	/protein_id="BAB53269.1"		595	ATGATCAGACCTTCTTGAACATGGCGCCCGCGCGATGACGCTGCAACAAAGTCCCGTTG	654
	/translation="MDAKALLAORGFHGRPMHSRHRCDRRRLQRLWAAAEAY AGGRLVGEINQAMIPWGRHIAARQART"		337762	TTCATCACCAACCATCTTCAACATCGCGCTCCGGCGATGACGCTGCACAAAGATCCCGCTG	337821
CDS	/db_xref="GI:14026674"		655	TTCTGCTGCTGCTGCTTATCACGGCTTGGCTGATCCTGCTGGCGCTTCGGGTCTCGCT	714
	/codon_start=1		337822	TTGGCTTGGTGGTGGTCAACCGCTTCTCTGTTGCTGCTCTCTGCTGCTGCTGCTGCTG	337881
gene	/product="GMC-type oxidoreductase"		715	GGTGAATCAACATGCTGCTGACCGACCGTAACTTCGCGACGACCTTCTCAATCTCTGCT	774
	/protein_id="BAB53270.1"		337882	GGCGGCATCACATGCTGCTCACCGACCGCAATTTTCGGCACCACTTCTTCGCCCGCAT	337941
CDS	/translation="MSETSFGDYDVIVAGSAGCVLANLSADPNRVLLLEAGGS RVHWDPIDPGLFCVGNPTDMKTEPEAGLNGSLNYPGKVLGGSSVNGMLYM RGAAADYDQRQAGNWDVLPYPRKSEDTGKSKSMHSGGSENRVERORLNWPI LDAPRAELGLPRTGDNTEGSEFSEVNRUGVRWNTSKAPFLAARQNLRV VTAEVERDFDGRVLTFRQDRLRVSAGQTLAAGAINSPKILSLSGVSPDL LSSIGVDRHALPGVGNLQHLQIRTVKVAANATLNERYNHLISRAAMGVEYALKR LSSIGVDRHALPGVGNLQHLQIRTVKVAANATLNERYNHLISRAAMGVEYALKR SRTVHMSADPEATKIRPNLSTAGDSVAVASLRHVRSLMKRAVARAPBEMLP GTYDSDDEDILRKAGDIGTIPHPVCTCKMSDSMAVVDRLRVHGLAGLRVVDASIM PTIVSGNTNSPVVIAEKAAEMILLDPRR"		775	GGCGCGGTGACCGATCTCTGACCAACATCTCTGTTGTTCTTTGGGACCCCGGAGTG	834
	/db_xref="GI:14026674"		337942	GGCGCGGTGACCGATCTCTGACCAACCTGTTCTGGTTCCTCGGTCTATCCGGAAGTG	338001
gene	/protein_id="BAB53270.1"		835	TACATCATCATTTGCCCGGCTTTGGCATCATCAGCCATCTCGTGTGACCTTCTCGAAA	894
	/codon_start=1		338002	TACATCTGATCTCTGCCGGCTTCCGGCATCATCAGCCATCTCTCGATCTCTCGAAG	338061
CDS	/transl_table=11		895	AAGCGGTCTTCGGTTACCTGCGCATGGTCTATGCAATGGTGGCAATCGTGTCTGGGC	954
	/product="GMC-type oxidoreductase"		338062	AAGCGGTCTTCGGTTATCTCGGCATGGCTTACGCCATGTCGCGCATCGGCGCGCTGCGC	338121
gene	/score="719.6; DB 1; Length 346510;"		955	TTTGTCTCTGGCGCACCATGTACACGTTTGGTATGCTGACCGACCAAGCAATCTCTAC	1014
	/best_local_similarity="65.7%; Pred. No. 2.5e-139;"		338122	TTCTGCTCTGGCGCACCATGTATACGACCGCGCTGCTCGCTCGACACGACGCGCTAC	338181
CDS	/mismatches="0; Mismatches 564; Indels 9; Gaps 3;"		1015	TTCTGCTGGCCACCATGGTGTATCGCGTGGCGACCGCATTAAGATCTTCTCGTGATC	1074
	/transl_table=11		338182	TTCTGCTGGCCACCATGGTGTATCGCGTGGCGACCGCATTAAGATCTTCTCGTGATC	338241
gene	/db_xref="GI:14026674"		1075	GCCACGATGGGCGCGCTGCTGTTGAGTTTCAAAATCGCCGATGCTCTGGGCTTTTGGCTTT	1134
	/codon_start=1		338242	GCGACGATGGGCGCGCTGCTGTTGAGTTTCAAAATCGCCGATGCTCTGGGCTTTTGGCTTT	338301
CDS	/transl_table=11		1135	ATGTTCTGTTTACCGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	1194
	/product="GMC-type oxidoreductase"		338302	ATCTTCTGTTTACCATCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	338361
gene	/db_xref="GI:14026674"		1195	CGTGCATATCAGCACCTTATAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	1254
	/codon_start=1		338362	CGTGCATATCAGCACCTTATAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	338421
CDS	/transl_table=11		1255	GCGATCTTTCGGATCTTTCGGCGGTGATCTACTTTTACATGCCGAAGTTCTCGGCGCGCT	1314
	/product="GMC-type oxidoreductase"		338422	GCGGTGTTCGCTCTTTCGGCGGTGATCTACTTTTACATGCCGAAGTTCTCGGCGCGCT	338481
gene	/db_xref="GI:14026674"		1315	TTCCCGAATGGGTGCGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	1374
	/codon_start=1		338482	TATTCGCCCGCTTCGCGCAACACCCACTTCTGGGTGACCTTTTCGGGTGACCTGATC	338541
CDS	/transl_table=11		1375	TTCTTCCCGCAGCATTTCTTGGGACGCTCAGGGTATCGCGCGCGCTTACATCGACTATCC	1434
	/product="GMC-type oxidoreductase"		338542	TTCTTCCCGCAGCATTTCTTGGGACGCTCAGGGTATCGCGCGCGCTTACATCGACTATCC	338601

```

QY 1435 GAAGCTTCGGCTGTGGAACAAAGTCTCTCTATGGTGGTCTCTCTGGCTTCGGCTCG 1494
Db 338602 GACCGGTTGCGCGCTGGAAGTCTCTCTATGGTGGTCTCTCTGGCTTCGGCTCGCGC 338661
QY 1495 TTCTGTTCTTCATCGTGTATCTTTGTTCTATACGCTGGTCTCTGGCGCGCGAGACCGCT 1554
Db 338662 GTGGCGATCTTCTCTACGGCGGTCTCGAAGCTTCCAGAGAGCGCATCG---CCGCT 338718
QY 1555 CCGAACCGTGGCGGAGATTCGCGGATACGCTGAGATGAGCGTGCATCAGCGCTCGC 1614
Db 338719 GCCAACCGATGGCGGAGGCGCGCACACGCTGGAATGGCAGTGGCTTCGGCGCGCGC 338778
QY 1615 GCCCACACGTTTCGAAACGCTGCCCAAGCGTTCGGAATCGGACATGGGACAAAGCATCC 1664
Db 338779 TTCCACCAGTGGAGCAGTTGCCGAAGATCAAGTAGCGCGCTCCCGCCC 338828

RESULT 13
NWCOXABC 4380 bp DNA linear BCT 27-JUL-2001
LOCUS N.winogradskyi DNA for coxA, coxB and coxC genes.
DEFINITION X89566
ACCESSION X89566
VERSION X89566.1 GI:1177572
KEYWORDS coxA gene; coxB gene; coxC gene; cytochrome c oxidase subunit I;
cytochrome c oxidase subunit II; Heme O synthase.
SOURCE Nitro bacter winogradskyi
ORGANISM Nitro bacter winogradskyi
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Bradyrhizobiaceae; Nitro bacter.
REFERENCE 1
AUTHORS Berben, G.
TITLE Nitro bacter winogradskyi cytochrome c oxidase genes are organized
in a repeated gene cluster
JOURNAL Antonie Van Leeuwenhoek 69 (4), 305-315 (1996)
MEDLINE 96433441
PUBMED 8836428
REFERENCE 2 (bases 1 to 4380)
AUTHORS Berben, G.P.R.
TITLE Direct Submission
JOURNAL Submitted (12-JUL-1995) G.P.R. Berben, Centre de Recherches
Agronomiques de l'Etat, Station de Chimie/Physique agric, Chaussee
de Wavre, 115, B-5030 Gembloux, BELGIUM
FEATURES
source
1. .4380
/organism="Nitro bacter winogradskyi"
/mol_type="genomic DNA"
/strain="ATCC 14123"
/db_xref="taxon:913"
/clone="plasmid Nw24.92"
/clone_lib="library of G.Berben"
727..1584
/gene="coxA"
727..1584
/gene="coxA"
727..1584
/gene="coxA"
/EC_number="1.9.3.1"
/codon_start=1
/transl_table=11
/product="cytochrome-c oxidase subunit II"
/protein_id="CAA61743.1"
/db_xref="GI:1177573"
/db_xref="GOA:Q59630"
/db_xref="TrEMBL:Q59630"
/translation="MMMSGQYRRRFRPLGLAFGGAFAAGAAFAADEIGQPHAW
GFDLQQVTPMSEIARMNGLLVITLITLVALLVAVYFRNARNPVSKTHN
TMIEVAWITIFVLIVGIAVPSRLIFQDLPLKADLTATKATQKQWYTYTPDNKIF
EPDLSLQDKKRLAVDNELVYVNVKVRVQVIGADVIHSFVPSFGMKVDVAPGLR
NETWFKATKGTGYGCGSELGRDHAFMPIAVRVVSEQFAAWVETAKKKFASNPANA
VASAAGVQW"
727..828
/sig_peptide
/gene="coxA"
1682..3301
/gene="coxA"
1682..3301
/gene="coxA"
1682..3301

```

```

/gene="coxB"
/EC_number="1.9.3.1"
/codon_start=1
/transl_table=11
/product="cytochrome-c oxidase subunit I"
/protein_id="CAA61744.1"
/db_xref="GI:1177574"
/db_xref="GOA:Q59631"
/db_xref="TrEMBL:Q59631"
/translation="MATAHADHDDHGDHHAHFTGWRKFAYSTNHNKIDIGTWYLVFA
IVAGITGLMSMGIRAEMLYPGIQVDPARTYNTVFNHGLINI FFMVMPMIGGFGN
WFVPLMIGAPDMAPFRMNNISWLLPFAAFALLISFMBEGEPQACAGTGMLYAPLS
TSGHPGPAVDPSLSLHLAGAGSILGAINFITTFNMRAPGMTKMPLEFVMSILVTV
ELLALPVLGAGAITWLTDRNPGTFRADGGDVLFOHLFWPFGPEVYVILLPG
FGMLSHVSTFSKPIFGILMAYANVAIGGIGVFWAHMYTVGNVSAQAYFVAAT
MWIAVPTGVKVFSDIATMMGSLFEPMLFAIFGLFTVGVTGCVLVANAGVDRVM
QETVYIAHPHYMGIAAFAIFSGYIYWFPGYMGYNETTIGKHFWMFPGANILF
PQHFGLQGMRRMYDYPDAFAGWNEISSYGAFTGEGVILIFLYGLVDAFARKQAA
DNWEGGATLTWLTSSPPPHQFSTLPKVQ"
3314..4303
/gene="coxC"
3314..4303
/gene="coxC"
/codon_start=1
/transl_table=11
/product="heme O synthase"
/protein_id="CAA61745.1"
/db_xref="GI:1177575"
/db_xref="GOA:Q51322"
/db_xref="TrEMBL:Q51322"
/translation="MCAASCAAHPTKAGSIUSVVDKAVELPPRISEAGVADYFALL
KPMVSLVIFTALVNIAPGHVHPVILICIAVAGAGSALNMALEDIDVIMS
RTNRPIPRITRGEAMGFTLSFFSVMTGLTVNAGGLAFTIFFYVVIYTMG
LKRTAQNIIVIGAGALPVAWAAATGSLVLELLIIFFMTPHFWALALFRS
DDYRAGVPMPLPVWAGPDATRLQILLYTVLVAIAAAPPLGYFDWYGVTSLLIAG
MLVLATEVYRHRTGSQALRATRLRPAFSIFYLFAVLVLLDVAKAVAPLIW"

ORIGIN
Query Match 42.4%; Score 709.2; DB 1; Length 4380;
Best Local Similarity 70.5%; Pred. No. 4.8e-137;
Matches 977; Conservative 0; Mismatches 403; Indels 6; Gaps 2;
QY 259 AACGTCATGTTACCTACCATGTTCTTCTGATGATGTTCTTGTGGGTATCCCGCATG 318
Db 1910 AACGTTTTCGACCAATCACGCGCTGATCATGATCTTCTCATGATGATGCTGGATG 1969
QY 319 TTGGTGGTTTGGTAACATCTGATGCGCGTCGAAATCGGCGCTCGGATATGGCTTC 378
Db 1970 ATGCGCGCTTCGCGCAACTGTTCTGCGCGCTGATGATCGGCGCGCGCGCATGGCTTC 2029
QY 379 CCGCGTATGAACAACCTGTCGTTCTGGCTGTTCTTCCGCTACCGCGTACCGCGTGGCT 438
Db 2030 CCGCGATGAACAACATCTGTTCTGGCTGCTCGCGCGCGCTTTCGCTGCTGATC 2089
QY 439 TCGCTGTTCCACCGCGCGTGAAGTCACTGCGGTTCGGGCTGGGTGGTGGTTCGTATC 498
Db 2090 TCTCGCTTCATGAGGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGTAT 2149
QY 499 CCGCGCTGTCGACCGCGCG---AAGTGGCTATTTCGATGAGCTCGCGATTTTCGCGT 555
Db 2150 GCSCCGCTGTCGACGAGCGGTGATCCGGTCCGCGCTGAGCTTTCAGCATCTGTCGATC 2209
QY 556 CATTGTCGGTGGCTTCCTGATCATGGCGGGAACAAATGATCAGCACTTCTTTGAC 615
Db 2210 CACTTTCGGCGCGCTTCCTGATCTTCGCGGCGATCACTTCACTCACCAGATCTTCAAC 2269
QY 616 ATGCGCGCGCGCGCATGACGCTGCAAAAAGTGCGCTTCTTCTCGTGGTGGATCTTTATC 675
Db 2270 ATGCGCGCGCTGCGATGACCAAGATGCGCGCTTCTTCTGCTGGTGGATCTTGGTC 2329
QY 676 ACGGCTTGGCTGATCTGCTGGCGCTTCGCGGTTCGCTGGTGGTGGTGGTGGTGGTGG 735
Db 2330 ACGGCTTGGCTGCTGCTGGCGCTTCGCGGTTCGCGCGCGCGCGATCAGATGCTGGTGG 2389

```


Qy	607	TTCTTGAACATGCGCGCCCGCGCATGACGCTGCACAAAGTGCCTGTGTTCTCTGTGGTGC	666
Db	763	ATCTTTCAACATGCGCGCGCGCGGCATGACCCCTGSCAACAGATGCCGCTGTTTCGTGTGGTGC	822
Qy	667	ATCTTTATCACGGGTTTGGCTGTATCTGTGTGGCGCTCCGGTTCCTGGCTGTGGTGCACATCACC	726
Db	823	ATCTGTGTGACCGTGTTTCCTGCTGCTGTGTTGCGTCCGCGTGCCTCGCGGTGCGATCACC	882
Qy	727	ATGCTGTGTGACCGACCGTAACTTCGCGACGACCTTCTCAATCTGCTGCGCGCGGTGCAC	786
Db	883	ATGCTGTCTACCGACCGCAATTTTCGCGACGACCTTCTTCGCGCGCGACGCGCGCGCGGCAC	942
Qy	787	CCGATTCTGTACCAACACATCTGTGGTTCCTTTTGGGCACCCGCGAAGTGTATCATCATTT	846
Db	943	CCCGTGTGTTCACGACCATCTGTCTGGTTCCTCGGCCATCCGAGGTGTATACCTCTGATC	1002
Qy	847	CTGCGCGGCTTTGGCATCATCAGCCATCTGCTGTCGACCTCTTCGAAAAAGCGCGTCTTC	906
Db	1003	CTGCGCGGCTTCGCGCATGATCAGCCAGATCTGCTCCACCTTCTCGCGCAAGCCCGTGTTC	1062
Qy	907	GTTTACCTGCGGATGGTCTATGCAATGGTGGCAATCGGTTCCTGGGCTTTGTGCTGTG	966
Db	1063	GCTATCTCGGCATGGCTACGGCATGGTTCGGCATCGCGCATCGGCTCTGTTGTTGGT	1122
Qy	967	CGCGACACATGTACACCGTGTGGTATGCTGCTGACCCAGCAATCCTACTTCATCTGCTGGCC	1026
Db	1123	CGCGACACATGTACACCGTGTGGCATGCTGCTCGCGACGCGAGGCTATTTCTGTGCGGCG	1182
Qy	1027	ACCATGTGTATCGCGTTCGCGACCGGCATTAAGAATCTTCTCTGTGGATGCGCAAGATGTGG	1086
Db	1183	ACGATGTGTATCGCGTTCGCGACCGGCTGGAAGATCTTCTGTGGATCGCGACGATGTGG	1242
Qy	1087	GCGCGCTCGTGTAGTTTCAAAATCGCGCATGCTCTGGGCTTTTGGCTTTATGTTCTCTGTTC	1146
Db	1243	GCGCGCTCGATCGAATCTTCGCGCGCGCATGATCTGGGCGTGGGCTCATCTTCTCTGTTC	1302
Qy	1147	ACCGTGGGTGTGTGACCGGTATCGTGTGCGCCCAAGCGGGTGTGGACCGTGCAATACAC	1206
Db	1303	ACGGTGTGTGTGACCGCGCTGCTGTGCGGGAACCGCGGCTGACCGCGTGTCTCCAG	1362
Qy	1207	GACACCTATTACGTTGGTGGCGCACTTCCATTATGTGATGTGCTGTGGTGGATCTTTGCG	1266
Db	1363	GAGACCTTACTAGTGTGCGCGCACTTCCACTAGTGTGCTGCTCGGTGCGATGTTCTCGG	1422
Qy	1267	ATCTTCGCGGTATCTACTTTTACATGCGGAAGTTCTCGGCGCGGCTTCCCGGAATGG	1326
Db	1423	ATCTTCGCGGCTGGTACTACTGGTTCCGGAAGATGACGGGTCATGTATACACGAGACG	1482
Qy	1327	GCTCGAAGCTGCACCTCTGGACCTTTCATCGGTGGGACGTCAGTGTCTTCTCCCGCAG	1386
Db	1483	CTGCGGAAGGCGCACTTCTGGGTCACTTCACTCGGGGTGAACCTGTGTCTTCTCCCGCAG	1542
Qy	1387	CACCTCTCGGAAGCTCAGGGTATCCGCGCGTTCATATGACTATCCGAGAGCTTCGCG	1446
Db	1543	CACCTCTCGGCTGTGCGGCATGCGCGCGCTACGTCGACTATCCGACGCGTTCGCC	1602
Qy	1447	CTGTGGAAACAAGTCTCGTCTATGTGTGCGTTCCTCGGCTTCGCTCGTCTCTGTTCTTC	1506
Db	1603	GGCTGGAACTGTGCTGCTCGTGGGCTCTTACATTTTCGGGTTTCGGGCTCTGATCTTC	1562
Qy	1507	ATGCTGTATCTTTGTATACGCTGGTGTGTGCGCGCGAGACCGCTCGAACCCTGTGG	1566
Db	1663	CTCTATTTCGCTGTATCGATGC---CTTTGGGAAAAAGTGCCTGGTGCACATATCCGTGG	1719
Qy	1567	GGCGAATTCGCGCATACGCTGGGAATGGACGCTGGCATCACCGCTCTCGGCGCCACCGTTC	1626
Db	1720	GGTCCGCGCGACACGCTGGAGATGGAAGCTGCGCTTCGCGCGCCCTTCCATCAGTTC	1779
Qy	1627	GAACGCTGCGCC	1638
Db	1780	GAAGTGTGCGCC	1791

RESULT 15	BUCOXA	2000 bp	DNA	linear	BCT 07-NOV-1991
LOCUS	B-japonicum coxa gene for subunit I of cytochrome c oxidase				
DEFINITION	(SC_number=1.9.3.1).				
ACCESSION	X54800				
VERSION	X54800.1	GI:39502			
KEYWORDS	cytochrome c oxidase.				
SOURCE	Bradyrhizobium japonicum				
ORGANISM	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Bradyrhizobiaceae; Bradyrhizobium.				
REFERENCE	1 (bases 1 to 2000)				
AUTHORS	Bott,M., Bolliger,M. and Hennecke,H.				
TITLE	Genetic analysis of the cytochrome c-a3 branch of the Bradyrhizobium japonicum respiratory chain				
JOURNAL	Mol. Microbiol. 4 (12), 2147-2157 (1990)				
MEDLINE	92211625				
PUBMED	1965217				
REFERENCE	2 (bases 1 to 2000)				
AUTHORS	Bott,M.				
TITLE	Direct Submission				
JOURNAL	Submitted (10-OCT-1990) Bott M., Eidgenossische Technische Hochschule, Mikrobiologisches Institut, Schmelzbergstrasse 7, CH-8092 Zurich, Switzerland				
FEATURES	Location/Qualifiers				
source	1. .2000				
	/organism="Bradyrhizobium japonicum"				
	/mol_type="genomic DNA"				
	/strain="USDA311b110spc4"				
	/db_xref="taxon:375"				
	/clone="PRJ3404"				
gene	75. .1700				
	/gene="coxA"				
CDS	75. .1700				
	/genes="coxA"				
	/codon_start=1				
	/transl_table=11				
	/product="cytochrome-c oxidase"				
	/protein_id="CAA38570.1"				
	/db_xref="GI:39503"				
	/db_xref="GOA:P31833"				
	/db_xref="Swiss-Prot:P31833"				
	/translation="WATSAARHGDAQHQHGHDEHAHGTGRRVYVSTNKHDKIGTWYLL FAVIVGIAGMSAIRAEMLPGVQIFHETHTYNTVSTHGLIMIFFWMPANIGGFF GWNVFLMTGADPAMFPRMNNISPMJLPASFGLLMSTFVEFGGANGAGAGTWMTVP LSSSGHPAVDPALSLHLGASSILGANFITTENRAPQMTLHKPELFWNSILV TFFLLLSLPVLAGAITMLLTDRNFGTTFAPDGGDGVLPQHLFWPFGHGVYIIL PPGMTSQVSTFSKPPVGLGWAYMAVIGGIFVVAWHMYTVGMSATQAVFA ATNVIAPGVKTIKSWLITMGSGSTEFRAPMIYAVGFILFTYGGVTGVVLNAGVDR VIGETYYVAHHYVLSLGVAFAPAGWYWFPGTGYMNEINLAKAHFWYTFIGNL VFFPQFLGSGMPRYVDYDFAPAGWNLVSSVGYSISGFGVLIPLYCVIDAPAKKVP AGNPNMGAGATLEWTLPSPPPPHQFVLPVQ"				
mat_peptide	75. .1697				
	/gene="coxA"				
	/product="cytochrome-c oxidase"				
CDS	1854. .>2000				
	/note="unnamed protein product"				
	/codon_start=1				
	/transl_table=11				
	/protein_id="CAA38571.1"				
	/db_xref="GI:39504"				
	/db_xref="GOA:Q99253"				
	/db_xref="TrEMBL:Q99253"				
	/translation="NSLVIFTLVGMAMAPQHPPHVLAIITSLLCIAVGAGASGALNNA LEGDI"				
mat_peptide	1854. .>2000				
	/product="unnamed"				
ORIGIN					
Query Match	42.3%	Score 708.8;	DB 1;	Length 2000;	
Best Local Similarity	70.3%;	Pred. No. 6.1e-137;			
Matches 979;	Conservative	0;	Mismatches 407;	Indels 6;	Gaps 2;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 22, 2004, 12:20:45 ; Search time 996 Seconds

(without alignments)
8822.822 Million cell updates/sec

Title: US-09-712-768A-1

Perfect score: 1674
Sequence: 1 atggcagacccgccattca.....acaagcatccctcgcaactaa 1674

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 segs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1674	100.0	1674	4 AAA91489	Aaa91489 G. oxydan
2	515.6	30.8	1593	8 ABT14590	Abt14590 Pseudomon
3	503.6	30.1	1359	11 ABD02101	Abd02101 Pseudomon
4	503.6	30.1	1587	11 ABD02072	Abd02072 Pseudomon
5	489	29.2	1278	11 ABD02134	Abd02134 Pseudomon
6	408.4	24.4	16568	8 ACC45136	Acc45136 Human mit
7	406.2	24.3	1629	8 ACA37676	ACA37676 Prokaryot
8	405	24.2	1555	5 AAS75523	Aas75523 DNA encod
9	398.8	23.8	16559	10 ADD42976	Add42976 Human mit
10	398.8	23.8	16561	10 ADD42971	Add42971 Human mit
11	398.8	23.8	16568	10 ADD43240	Add43240 Human mit
12	398.8	23.8	16569	10 ADD43266	Add43266 Human mit
13	398.8	23.8	16569	10 ADD43327	Add43327 Human mit
14	398.8	23.8	16569	10 ADD43295	Add43295 Human mit
15	398.8	23.8	16569	10 ADD43302	Add43302 Human mit
16	398.8	23.8	16569	10 ADD43334	Add43334 Human mit
17	398.8	23.8	16569	10 ADD43355	Add43355 Human mit
18	398.8	23.8	16569	10 ADD43350	Add43350 Human mit
19	398.8	23.8	16569	10 ADD43007	Add43007 Human mit
20	398.8	23.8	16569	10 ADD43300	Add43300 Human mit
21	398.8	23.8	16570	10 ADD42959	Add42959 Human mit

ALIGNMENTS

RESULT 1

AAA91489

ID AAA91489 standard; DNA; 1674 BP.

XX AC AAA91489;

XX DT 06-AUG-2001 (first entry)

XX DE G. oxydans cytochrome C oxidase (COI) coding sequence.

XX KW Cytochrome C oxidase; COI; COII; COIII; COI complex; D-sorbitol;
oxidative fermentation; electron transfer; respiratory chain; L-sorbose;
2-Keto-L-gluconic acid; 2KGA production; aldehyde production;
carboxylic acid production; ketone production; ds.
XX OS Gluconobacter oxydans.

XX FH Key Location/Qualifiers

XX FT CDS 1..1674

XX FT /*tag= a

XX FT /product= "COI"

XX PN EPI103603-A2.

XX PD 30-MAY-2001.

XX PF 14-NOV-2000; 2000EP-00124785.

XX PR 17-NOV-1999; 99EP-00122842.

XX PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX PI Asakura A, Hoshino T, Shinjoh M;

XX DR WPI; 2001-357953/38.

XX DR P-PSDB; AAY97750.

XX PT New cytochrome c oxidase complex having cytochrome c oxidase activity
from Gluconobacter oxydans DSM 4025, useful in mediating electron
transfer in respiratory chain or producing 2-keto-L-gulonic acid from L-
sorbose or D-sorbitol.

XX PS Claim 13; Page 16-20; 42pp; English.

XX XX

CC This sequence encodes the Glucanobacter oxydans cytochrome C oxidase
 CC (COI) of the invention. The COI complex is useful in improving oxidative
 CC fermentation and is an essential component mediating electron transfer in
 CC the respiratory chain. The recombinant microorganism and the cytochrome c
 CC oxidase may be used in the genetic preparation of a recombinant COI
 CC complex and in the production of 2-keto-L-gluconic acid (2KGA) from L-
 CC sorbose or D-sorbitol in a culture medium. The COI is also useful as a
 CC terminal oxidase, oxidising cytochrome c, an electron acceptor from an
 CC enzyme belonging to dehydrogenase for the production of aldehydes,
 CC carboxylic acids and ketones from alcohols and aldehydes, especially the
 CC production of 2KGA from L-sorbose or D-sorbitol
 XX
 SQ Sequence 1674 BP; 270 A; 493 C; 468 G; 443 T; 0 U; 0 Other;

Query Match 100.0%; Score 1674; DB 4; Length 1674;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCAGACGCGCCGATTCACGGCCATGACACCATGAGAGCAAGGCTTCTTTCACGCGC 60
 DB 1 ATGGCAGACGCGCCGATTCACGGCCATGACACCATGAGAGCAAGGCTTCTTTCACGCGC 60

QY 61 TGGTTTCATGTCGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 120
 DB 61 TGGTTTCATGTCGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 120

QY 121 GTTGGTTTCATGTCGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 180
 DB 121 GTTGGTTTCATGTCGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 180

QY 181 GTTCAGTACATGTCGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 240
 DB 181 GTTCAGTACATGTCGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 240

QY 241 GCGAAGGACACCTGTGGAACGTCATGTTTACCTACCATGATGTTTCTGATGATGTTTCTTT 300
 DB 241 GCGAAGGACACCTGTGGAACGTCATGTTTACCTACCATGATGTTTCTGATGATGTTTCTTT 300

QY 301 GTGGGTATCCCGCATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 360
 DB 301 GTGGGTATCCCGCATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 360

QY 361 GCTCCGGATATGGCTTCCCGCATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 420
 DB 361 GCTCCGGATATGGCTTCCCGCATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 420

QY 421 ACCGGATGGGGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 480
 DB 421 ACCGGATGGGGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 480

QY 481 GTTGGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 540
 DB 481 GTTGGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 540

QY 541 GCGATTTTCGGGTTTCACTTTCGGGTTTGGGTTTGGGTTTGGGTTTGGGTTTGGGTTTGGGTTTGGG 600
 DB 541 GCGATTTTCGGGTTTCACTTTCGGGTTTGGGTTTGGGTTTGGGTTTGGGTTTGGGTTTGGGTTTGGG 600

QY 601 ACGACCTTCTTGAACATCGCGGCCCCCGGCGATGACGCTGACCAAGTGGCGTTGTTCTCG 660
 DB 601 ACGACCTTCTTGAACATCGCGGCCCCCGGCGATGACGCTGACCAAGTGGCGTTGTTCTCG 660

QY 661 TGGTGGATCTTATCAGCGCTTGGGTGATCTGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 720
 DB 661 TGGTGGATCTTATCAGCGCTTGGGTGATCTGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 720

QY 721 ATCACCATGCTGCTGACCGACCGTAACCTTCGGACGACGCTTCTTCAATCTGCTGGCGGC 780
 DB 721 ATCACCATGCTGCTGACCGACCGTAACCTTCGGACGACGCTTCTTCAATCTGCTGGCGGC 780

QY 781 GGTGACCGGATCTGTACCAACATCTGTTGTTCTTGGGACCGCGGAAGTGTACATC 840
 DB 781 GGTGACCGGATCTGTACCAACATCTGTTGTTCTTGGGACCGCGGAAGTGTACATC 840

QY 841 ATCATTTGCCCGCTTTGGCATCATCAGCATATGCTGTGCGACCTTCTCGAAAAAGCGG 900
 DB 841 ATCATTTGCCCGCTTTGGCATCATCAGCATATGCTGTGCGACCTTCTCGAAAAAGCGG 900

QY 901 GTCTTCGTTACCTTCGCGATGCTTATGCAATGCTGCAATCGTGTCTTGGGCTTTGTC 960
 DB 901 GTCTTCGTTACCTTCGCGATGCTTATGCAATGCTGCAATCGTGTCTTGGGCTTTGTC 960

QY 961 GTCTGGGCGACCAACATGATGATACACGCTTGGTATGTGCTGACCGCAATCTTCTATCTTATG 1020
 DB 961 GTCTGGGCGACCAACATGATGATACACGCTTGGTATGTGCTGACCGCAATCTTCTATCTTATG 1020

QY 1021 CTGGCCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
 DB 1021 CTGGCCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080

QY 1081 ATGTGGGCGCTCGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
 DB 1081 ATGTGGGCGCTCGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140

QY 1141 CTGTTCACCGTGGGCTGTGACCGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
 DB 1141 CTGTTCACCGTGGGCTGTGACCGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200

QY 1201 TATCAGCAGACCTATTACGTTGGTGGGCACTTCAATATGATGATGCTGCTGCTGCTGCTGCTGCTGCTG 1260
 DB 1201 TATCAGCAGACCTATTACGTTGGTGGGCACTTCAATATGATGATGCTGCTGCTGCTGCTGCTGCTGCTG 1260

QY 1261 TTTGGGATCTTCGCGGCTATCTATCTTATCATCCGAGTTCGCGGCGCGCTTCCCG 1320
 DB 1261 TTTGGGATCTTCGCGGCTATCTATCTTATCATCCGAGTTCGCGGCGCGCTTCCCG 1320

QY 1321 GAATGGGCTGCAAAAGCTGCACTTCTGGAACCTTCTTATCGGTGCGAACGTCAGTTCTTC 1380
 DB 1321 GAATGGGCTGCAAAAGCTGCACTTCTGGAACCTTCTTATCGGTGCGAACGTCAGTTCTTC 1380

QY 1381 CCGCAGACACTTCTCGGACGTCAGGATATCCCGCGCGTATCATCGACTATCCCGAAGCC 1440
 DB 1381 CCGCAGACACTTCTCGGACGTCAGGATATCCCGCGCGTATCATCGACTATCCCGAAGCC 1440

QY 1441 TTCGCGTGTGGAAACAAAGTCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
 DB 1441 TTCGCGTGTGGAAACAAAGTCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500

QY 1501 TTTCTCATCGTGTGCTTGTGCTATACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
 DB 1501 TTTCTCATCGTGTGCTTGTGCTATACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560

QY 1561 CCGTGGGCGCAATTCGCGGATGACGCTGGAATGACGCTGCAATCCGCTCCGCGCCAC 1620
 DB 1561 CCGTGGGCGCAATTCGCGGATGACGCTGGAATGACGCTGCAATCCGCTCCGCGCCAC 1620

QY 1621 ACGTTCGAACGCTGCGCAAGCGCTGGAATGACGCTGGAACAGCATCCCTCGCACTAA 1674
 DB 1621 ACGTTCGAACGCTGCGCAAGCGCTGGAATGACGCTGGAACAGCATCCCTCGCACTAA 1674

RESULT 2

ABT14590

ID ABT14590 standard; DNA; 1593 BP.

XX ABT14590;

XX 27-FEB-2003 (first entry)

XX Pseudomonas aeruginosa biofilm formation-related gene #32.

XX Gene; ds; biofilm formation modulation; biofilm-associated disease;

XX cystic fibrosis; AIDS; middle ear infection; acne; periodontal disease;

XX catheter-associated infection; medical device-associated infection.

XX Pseudomonas aeruginosa.

OS

XX PN W0200285295-AZ.
 XX PD 31-OCT-2002.
 XX PF 19-APR-2002; 2002WO-US012532.
 XX PR 20-APR-2001; 2001US-0285190P.
 XX PR 24-OCT-2001; 2001US-0344142P.
 XX PA (IOWA) UNIV IOWA RES FOUND.
 XX PA (HARD) HARVARD COLLEGE.
 XX PI Whiteley M, Bangera MG, Lory S, Greenberg BP;
 XX DR WPI; 2003-075601/07.
 XX DR P-PSDB; ABJ18768.
 XX PT Identifying compound capable of modulating biofilm formation by
 XX PT bacteria/bacterial antibiotic resistance, useful for treatment of biofilm
 XX PS associated disease.
 XX PS Claim 1; Page 76; 154pp; English.
 XX CC The invention comprises a method for identifying a compound capable of
 XX CC modulating biofilm formation by bacteria. The method of the invention is
 XX CC useful for identifying a compound capable of modulating biofilm formation
 XX CC by bacteria or modulating bacterial antibiotic resistance. The method of
 XX CC the invention is also useful for diagnosing and treating a subject
 XX CC (especially an immunocompromised human) that is afflicted with a biofilm-
 XX CC associated disease or disorder, such as: cystic fibrosis; AIDS; middle
 XX CC ear infections; acne; periodontal disease; catheter-associated infections
 XX CC ; and medical device-associated infections. The present DNA sequence
 XX CC represents a gene that is used in the invention
 XX SQ Sequence 1593 BP; 238 A; 568 C; 451 G; 336 T; 0 U; 0 Other;
 Query Match 30.8%; Score 515.6; DB 8; Length 1593;
 Best Local Similarity 63.0%; Pred. No. 8.4e-128;
 Matches 854; Conservative 0; Mismatches 484; Indels 18; Gaps 3;
 QY 286 CTGATGATGTTCTTTGTGGGTAACCCCGATGTTTCGGTGTGTTGTAATCTGATG 345
 DB 247 CTGATCATGGTCTTCGGGCGGTGATGCGGGCTTCGTGGGCTGGCCAACTGGATGATC 306
 QY 346 CCGCTGCAATCGCGCTCCGGATGCGCTTCCCGGTATGAACAACTGCTGTTCTGG 405
 DB 307 CCGCTGATGATCGGGCGCGCGGACATGGCCCTCGCGGATGAACAACTTACGTTCTGG 366
 QY 406 CTGTTCAATGCGCGTACCGCATGGCGGTGGCTTCGCTGTTCCGACCGGGGGTGACGGT 465
 DB 367 CTGCTCCCGCGGCTTCGGCTGCTGTTGTCAGACACTTGTTCATGCCGGGGG----- 419
 QY 466 CAGCTGGGTTGGGCGTTGGTGGTTCTGTACCGCGCGCTTCGACCGCGGAAGCTGGC 525
 DB 420 -----CGGCCCCCAACTTCGGCTGGACCTTCTATGCGCGCTCTGACACCACTTCGCCCGCG 474
 QY 526 TATTCGATGCACTCCGATTTTCGGGTTCACTTGTTCGGGTGCTCTCCGATCATGGGC 585
 DB 475 CACAGCTGACCTTCTTCACTTCGCACTCCACCTGGCCGGGATCAGCTCGATCATGGGC 534
 QY 586 GCGATCAACATGATACAGACCTTCTTGAACATGGCGGCCCGCGGATGACGCTGCACAA 645
 DB 535 GCGATCAACGATGATCGCCACCACTCTCAACCTGCGCGCCCGGGCATGACCTGATGAAG 594
 QY 646 GTCCCGTGTTCGTGGTTCGATCTTATCAGCGTTGGCTGATCCTGCTGGGTGGCG 705
 DB 595 ATGCCGTGTTCGTGGACCTGGCTGATCACCCTGCTTCTTTGATCGGTGATGCT 654
 QY 706 GTTCTGGCTGGTCAATCAACATGCTGCTGACCGACCGTAACCTTCGGACGACCTTCTTC 765
 DB 655 GTGCTGGCGCGGTGTGACCATGATGCTGATGACATCCACTTCGGCACCAAGCTTCTTC 714

QY 766 AATCTCTGCTGGCGGGTGAACCCGATTCGTACCAACACATCCTCTGGTGTCTTTGGGCAC 825
 DB 715 ACGCGCGCGCGCGCGGACCCGGTGTCTTCCAGCAGCTGTTCTGGTCTTTCGGCCAC 774
 QY 826 CCGGAAGTGTACATCATCTCTGCCCGCTTTGGCATCATCAGCCATGTCTGTCTGACCC 885
 DB 775 CCGGAGGTGTACATCATGATCTGCGCCCTTCGGTGGGTGAGTGGCATCATCCCGACC 834
 QY 886 TTCTCGAAAAACCGGTCTTCGGTTACCTGCGGATGCTGTATGCAATGTTGGCAATCGGT 945
 DB 835 TTTCGCGCAAGCCGCTGTTCGGCTACACCTCGATGGTCTACGCCACCGCAGCATCGCC 894
 QY 946 GTTCTGGGCTTTGTCTCTGGCGCACCAATGTATACCGTTGGTATGTCTGTGACCCAG 1005
 DB 895 TTCTCTCTCTGTTGTTCTGGCGCACCAATGTATGTTCTGGTCTCGCATCCCGGTCA 954
 QY 1006 CAATCTCTCTGATGCTGCGCACCATGTGTGATGCGGTGCGGTCGCCACCGGCAATTAAGATCTTC 1065
 DB 955 GAGCTCTTCTTATGATGACCCACCATGCTGATCGCGGTGCCACCGGGGTGAAGGTGTTTC 1014
 QY 1066 TCGTGTGATCGCCACGATGTTGGGCGGCTCGGTGTAGTTCAAAATCGCCGATGCTCTGGGCG 1125
 DB 1015 AACTGGGTGACACCATGTTGGAGGGTTCGCTGACCTTCGAGACCGCGATGCTGTTCCGCC 1074
 QY 1126 TTTGGCTTTATGTTCTCTTCAACCGTGGTGTGTGACCGGTATCTGCTGGCCCAAGCG 1185
 DB 1075 GTGGCTTCTGTCATCTCTGTTTCAACATCGCGGGCTTCTCGGACTGATGTGGCGATCGCC 1134
 QY 1186 GGTCTGACCGGTGATATCAGACACCTATTACGTGGTGGCGCACTTCCATTATGTGATG 1245
 DB 1135 CCGGAGAGTTCAGTACACGACACCTACTTCTGGTGGTGGCCACTTCCACTACGTGCTG 1194
 QY 1246 TCGCTGGTGGCATCTTTGCGATCTTTCGGCGGTATCTATTTTACATGCGGAAAGTTCTCG 1305
 DB 1195 GTGCCCGGCGCATCTTGGCATCTTTCGCTCGGCTTACTTGGCTGGCGAAGTGGACC 1254
 QY 1306 GGCGCGCTTTCGGGAATGGGTGCAAGCTGCATCTTGGACCTTCTTCATCGGTGG 1365
 DB 1255 GGCCACATGTACGACGAGACCTTCGGAAGCTGCATCTTGGATGAGCTTTCATCGGCATG 1314
 QY 1366 AAGCTACGCTTCTTCCGACGACCTTCTCGGGACGTGAGGTAATGCGCGCCGTTACATC 1425
 DB 1315 AACTGCGGTCTTCCCGATGCACTTCTGCGGCTCGCGGCATGCGCGACGATCCCG 1374
 QY 1426 GACTATCCCGAAGCTTTCGGCTGTGGAAACAAAGTCTCTCTATGGTGGTCTTCTGGCC 1485
 DB 1375 GACTACAACTGCAAGTTCGCGGACTTCAACATGTTCTGTCGATCGCGGCTTTCATG 1431
 QY 1486 TTGCGCTCGTTCCTGTTCTTTCATCGTATCTTTGTCTATACGCTGTTGTCGCGCGCCG 1545
 DB 1432 TTGCGCACCACTGCTGTTCTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCT 1491
 QY 1546 GAGACCGCTCCGAACCGGTGGGGAATTCGCGGATACGCTGGAATGGAAGCTGCGCATCA 1605
 DB 1492 CCGGCCCTTCGAAGCCCTGGGACG---GGCGGAGGCGCTGGAGTGGAGCATCCCGCTCG 1548
 QY 1606 CCGCTCGGGCCACGATTCGAACCGCTGCCAAG 1641
 DB 1549 CCGCGCGCTTACCAACACTTCAGCACCCCGCCGAG 1584

RESULT 3

ABD02101

ID ABD02101 standard; DNA; 1359 BP.

XX

AC ABD02101;

XX

DT 29-JUL-2004 (first entry)

XX

DE Pseudomonas aeruginosa polynucleotide #705.

XX

KW Bacterial infection; gene; ds; Pseudomonas aeruginosa infection; antibacterial.

XX 29-JUL-2004 (first entry)
XX Pseudomonas aeruginosa polynucleotide #676.
XX Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
KW antibacterial.
XX Pseudomonas aeruginosa.
OS US6551795-B1.
XX PN 22-APR-2003.
XX PF 18-FEB-1999; 93US-00252991.
XX PR 18-FEB-1998; 98US-0074788P.
XX PR 27-JUL-1999; 98US-0094190P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
DR WPI; 2003-615309/58.
XX P-PSDB; ABO68501.
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX Disclosure; SEQ ID NO 676; 455pp; English.
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABD01397-
CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX SQ Sequence 1587 BP; 342 A; 458 C; 566 G; 221 T; 0 U; 0 Other;
Query Match 30.1%; Score 503.6; DB 11; Length 1587;
Best Local Similarity 63.5%; Pred. No. 1.4e-124;
Matches 841; Conservative 0; Mismatches 464; Indels 19; Gaps 4;
QY 319 TTCCGGTGGTTTGGTAACATATCGATGCGCGTCAATCGCGCTCCGGATATGG-CCTT 377
DB 1420 TTGCTCGGCTGGCCAACTGGATGATCCCGCTGATGATCGGGCGCGGACATGCCCT 1361
QY 378 CCCCGGTATGAACAACCTGCTGCTTCTGGTGTTCATTTGCCGATGCGCGTGGC 437
DB 1360 GCGCGGATGAACAATTTACGCTTCTGGCTGCTGCGCGCGGCTTCGGCTGCTGTCAG 1301
QY 438 TTGCTGTTCGACCCGGCGGTACCGTACGTGCTGGGTTGCGGCGTGGTTCGTGTA 497
DB 1300 CACCTGTTCATCCCGCGGCGGCGGCCCCAA-----CTTCGGGTGACCTTCTA 1253
QY 498 CCCGCGCTCGACCCCGCGGAAGCTGGCTATTCGATGACCTCGCGATTTTCGGGTTC 557
DB 1252 TGGCGCGTGTGACCACTTCGCGCGCACACGCTGACCTTCTTCATCTTCGCCATCCA 1193
QY 558 CTTGTGGGTGCTCTCTCGATCATGGCGCGATCAACATGATCAGACCTTCTTGACAT 617

DB 1192 CTTGGCGGGATCAGCTCGATCATGGCGCGGATCAACGATGATGCGCACCATCTCAACCT 1133
QY 618 GCGCGCCCCCGGCATGACGCTGCACAAAGTGGCGTGTCTCGTGTGATCTTTATCAC 677
DB 1132 GCGCGCCCCCGGCATGACGCTGCATGAGATGCGGCTGTCTGCTGACCTGGCTGATCAC 1073
QY 678 GGTTCGGCTGATCTCTGCTGCGCTGCGGCTTCTGGTGTGCAATCACCATGCTGCTGAC 737
DB 1072 CGGTTCTCTGCTGATCGCGTGTATGCGGCTGTGGCGGCGTGGTGTGATCATGATGAT 1013
QY 738 CGACCGTAATCTTCGGCACGACCTTCTCAATCTGCTGGCGGCGGTGACCCGATCTGTA 797
DB 1012 GGACATCCACTTCGGCACAGCTTCTTCAGCGCGCGCGCGCGGCGACCCGGTGTGT 953
QY 798 CCAACACATCTCTGTGTTCTTTGGGACCCCGGAAGTGTATCATCATCTTCTGCCCCGCTT 857
DB 952 CCAGCAGCTGTTCTGGTCTCTTCGCGCACCCGAGGTGTATCATCATGATCTGCGCGCTT 893
QY 858 TGGCATCATCAGCATGCTGCTGCGACCTTCTCGAAGAGCGGTCTTCTGGTTACCTGCC 917
DB 892 CGGTGGGTGAGTGCATCATCCGACCTTCGCGCGCAAGCCGCTGTTGGCTACACCTC 833
QY 918 GATGGTCTATGCAATGTTGGCAATCGGTGTTCTTGGGCTTTGTCTCTGGGCGCACCAT 977
DB 832 GATGGTCTACGCCACCCGCGACATCGCTTCTCTCTCTGCTGCTGCGCGCACCAT 773
QY 978 GTACACCGTTGGTATGTCGTGACCCGACGATCTTACTTCTATCTGCTGGCGCACCATGAT 1037
DB 772 GTTCGTGTCGGATCCCGTCAACCGCGAGCTGTTCTTTCATGACCGCACCATGCTGAT 713
QY 1038 CGGCGTCCGACCCGCGATTAAGATCTTCTGCTGGATGCGCCACGATGTGGCGCGCTCGGT 1097
DB 712 CGCGTCCGACCCGCGGTGAAGTGTTCACCTGGGTGACCCACCATGTGGGAGGTTGCT 653
QY 1098 TGAAGTCAAAATGCGCGATGCTCTGGGCTTTGGCTTTATGCTTCTGTTACCGTGGTGG 1157
DB 652 GACCTTCEAGACGCGGATGCTGTTGCGCGTGGGCTTCTGTCATCTGCTACCATTCGCGG 593
QY 1158 TGTGACCGGTATCGTGTGCGCCCAAGCGGTCTGGACCGTGCATATCACCACACCTATTA 1217
DB 592 CTTCTCGGACTGATGCTGGGATGCGCCCGCGGAGCTTCAGTACACGACACCTACTT 533
QY 1218 CGTGGTGGCGCACTTCCATTATGATGTCGCTGGGTGGGATCTTTCGATCTTCGCGCG 1277
DB 532 CGTGGTGGCGCACTTCCACTACGTCGTGTTGCGCGCGGATCTTCGCGCATCTTCGCTC 473
QY 1278 TATCTACTTTACATGCGGAGTTCGCGCGCGCTTTCGCGAATGGGCTGCAAGCT 1337
DB 472 GGCTTACTTCTGCTGCGGAGTGGACCGGCGCATGTACGACGAGACCTTCGCGAGCT 413
QY 1338 GCACCTTCTGGACCTTCTTTCATCGGTGCGAAGCTCACGTTCTTCCGCGACACTTCCTGGG 1397
DB 412 GCACCTTCTGGATGAGCTTCATCGGCATGAACCTGGGCTTCTTCCGATGACCTTCGTCG 353
QY 1398 ACCTCAGGTTATGCGCGCGTTCATGCTGACTATCCGAGGCTTCGCGGTGTGGACAA 1457
DB 352 CTTCCGCGGATGCGCGGACGCGATCCCGGACTACACCTCGAGTTCGCGCATCTCAACAT 293
QY 1458 AGTCTGCTCTATGCTGGTTCCTGCGCTTCTGCTGCTTCTTCTTCTGCTGATCTT 1517
DB 292 GGTCTGCTGCTGATCGCGGCTTCTG---TTGCGCACCAACCGCTGCTGTTCTGTTTCA 236
QY 1518 TGTCTATACGCTGCTGCTGCGCGCGGAGACCCGCTCGAACCCTGGGCGGATTCG 1577
DB 235 CGTCATCAAGTGCATCCGCGCGGCAAGCGCGCTTCCCAAGGCTTCGGACG---GCGC 179
QY 1578 CGATCGGTGAATGCGCTGCGCTGCGCTCACCGCTTCCGCGCGCACACCTTCGAAACGCTGCC 1637
DB 178 CGAGGCTGAGTGGAGCATCCCTCTCGCGCGGCGCTTACCAACCTTCAGCACCCCGCC 119
QY 1638 CAAG 1641
DB 118 CGAG 115

```
RESULT 5
ABD02134
ID ABD02134 standard; DNA; 1278 BP.
XX
AC ABD02134;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polynucleotide #738.
XX
KW Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
KW antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
FN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252391.
XX
PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094130P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI; 2003-615309/58.
XX DR P-PSDB; ABO68563.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 738; 455pp; English.
XX
CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABD01397-
CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 1278 BP; 182 A; 463 C; 361 G; 272 T; 0 U; 0 Other;
Query Match 29.2%; Score 489; DB 11; Length 1278;
Best Local Similarity 63.6%; Pred. No. 1.1e-120;
Matches 818; Conservative 0; Mismatches 450; Indels 19; Gaps 4;
QY 356 TCGCGCTCGGATATGG-CTTCCCGCTATGAACAACCTTCGTTCTGGCTGTTCATT 414
DB 1 TCGCGCGCGGACATGGCCCTCGCGGATGAACAATTTTCAGTTCTGGCTGTGCGG 60
QY 415 GCGGGTACCGGATGGCGTGGCTTCGCTGTCGACCGCGGGGTCACGGTCAGTGGGT 474
DB 61 GCGGCTTGGCGTGTGGTCACACCCCTGTTCATGCGCGGGCGGCCCCAA----- 113
QY 475 TCGGGCGTGGTGGTTCGTACCCCGCGCTGTGACCCGCGAGCTGGCTATTTCGATG 534
DB 114 -----CTTCGGCTGGACCTTCTATGGCGCGCTGTGACCACTTCGCCCCGACAGCGTG 168
```

QY	1615	GCCACACGTTGGAACCGTACCTACCTGGTATCTTGTGATGTTCTTTGTG	303
Db	1243	TACCAACACCTTCAGACACCCCGCCGAG	1269
XX	RESULT 6		
AC	ACC45136		
XX	ID	ACC45136 standard; DNA; 16568 BP.	
AC	ACC45136;		
XX	11-JUN-2003	(first entry)	
DT	XX	Human mitochondrial DNA sequence SEQ ID NO:1.	
XX	XX	Human; mitochondrial DNA; mtDNA; allele; diagnosis; haplotype;	
KW	XX	criminal investigation; forensic analysis; energetic imbalance;	
KW	XX	energy metabolism-related physiological condition; metabolic disease;	
KW	XX	abnormal energy metabolism; abnormal temperature regulation; diabetes;	
KW	XX	abnormal oxidative phosphorylation; abnormal electron transport; obesity;	
KW	XX	hypertension; cardiovascular disease; gene; ds.	
OS	XX	Homo sapiens.	
XX	XX		
XX	XX	Key	Location/Qualifiers
FT	XX	misc_feature	1..16568
FT	XX		*tag= a
FT	XX		/note= "Nucleotide alleles are given using the IUPAC
FT	XX		ambiguity code bases e.g. y for T or C"
XX	XX	WO2003018775-A2.	
XX	XX	06-MAR-2003.	
PD	XX		
XX	XX	30-AUG-2002; 2002WO-US028471.	
PF	XX		
XX	XX	30-AUG-2001; 2001US-0316333P.	
PR	XX	31-AUG-2001; 2001CA-02356536.	
PR	XX	13-MAY-2002; 2002US-0380546P.	
XX	XX		
XX	XX	(UYEM-) UNIV EMORY.	
PA	XX		
XX	XX	Wallace DC, Hosseini S, Mishmar D;	
PI	XX		
XX	XX	WPI; 2003-290064/28.	
DR	XX		
XX	XX	Diagnosing a haplogroup from a human, useful in criminal investigations	
PT	XX	and forensic analyses by providing a sample having mitochondrial nucleic	
PT	XX	acid, and identifying the presence of nucleotide allele/s diagnostic of a	
PT	XX	haplogroup.	
XX	XX		
PS	XX	Claim 23; Page 117-121; 126pp; English.	
XX	XX		
CC	XX	The present invention describes a method for diagnosing a haplogroup from	
CC	XX	a human comprising providing a sample comprising mitochondrial DNA	
CC	XX	(mtDNA) from the human, and identifying in the sample, the presence of at	
CC	XX	least one nucleotide allele diagnostic of a haplogroup. The method of	
CC	XX	diagnosing haplogroups of a human is useful in criminal investigations	
CC	XX	determining the presence or absence of nucleotide alleles. The alleles	
CC	XX	are useful for diagnosing predisposition to an energy metabolism-related	
CC	XX	physiological conditions such as energetic imbalance, metabolic disease,	
CC	XX	abnormal energy metabolism, abnormal temperature regulation, abnormal	
CC	XX	oxidative phosphorylation, abnormal electron transport, obesity,	
CC	XX	diabetes, hypertension, or cardiovascular disease. The present sequence	
CC	XX	represents a human mtDNA sequence which shows alleles in the sequence	
CC	XX	with IUPAC ambiguity code bases e.g. y for T or C	
XX	XX		
SQ	XX	Sequence 16568 BP; 4859 A; 4950 C; 1991 G; 3829 T; 0 U; 939 Other;	
XX	XX		
XX	XX	Query Match	24.4%; Score 408.4; DB 8; Length 16568;
XX	XX	Best Local Similarity	56.8%; Pred. No. 1.2e-98;
XX	XX	Matches	726; Conservative 32; Mismatches 508; Indels 12; Gaps 2;

QY	244	AACGGACACCTGTGGAAACGTCATGGTTACCTACCATGGTATCTTGTGATGTTCTTTGTG	303
Db	6050	AACGACACATCTACAAACGTYATCGTYACAGCCCATGCAATTTGTAATAATCTTTTTCATA	6109
QY	304	GGTATCCCGCAATTGTTCCGTTGGTATCTATCTGATGCCGTGCAAAATCGCGCT	363
Db	6110	GTAATACCCATCATAATCGAGGCTTTGGCACTGACTARTYCCCTAATAATATYGGVGC	6169
QY	364	CCGATATGGCCTTCCGGGTATGAACAACCTGCTGTTCTGGCTGTTTCATTTGGCGTACC	423
Db	6170	CCGATATGGCTTCCCGCATAAACAACATAGCTTCTGACTCTTACCTCTTACCTCTCTC	6229
QY	424	CGGATGGCGTGGCTTCCGCTGTTCCGACACCGGGGGGTGAGGTTGAGTGGTTCGGGGGTT	483
Db	6230	CTACTCTCTCGCATCTGCTATAGTRGAGGCGGAGCAGGAACAGTTGAACAGTC---	6286
QY	484	GGTGGGTTCTGACCCGCGCTGTGACCCGCGAAGCTGGCTATTCGATGGACCTCGG	543
Db	6287	-----TACCCTCCCTTAGCAGGGAATCTCTCCACCCCTGGARCCCTCGTAGACCTAAC	6340
QY	544	ATTTTCGGCTTCACTTGTCCGGTGGCTCTCTGATCATGGGCGCGATCAACATGATCAG	603
Db	6341	ATCTTCTCCTTACACCTAGCAGGTTCTCTATCTTAGGGGCCATCAATTTCAATCACA	6400
QY	604	ACCTTCTTGAACATGCGCGCCCGCGCAACGCTGCAACAAGTGGCTTCTCTGTTG	663
Db	6401	ACAATTATCAATATAAAACCCCTGCGATAACCAATACCAACCGCCCTCTTCTGTTGA	6460
QY	664	TGATCTTTATCAGGCTTGGCTGATCCTGCTGGCGTGGCGTCTGCTGGTGGTGAATC	723
Db	6461	TCGTCCTAATYACAGCAGTCCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	6520
QY	724	ACCATGCTGTGACCGACCGCTTAACCTCGGCACGACCTTCTTCAATCTGTGCGCGGT	783
Db	6521	ACTATCTACTACAGACCGCACTTAAACACACCTTCTTTCGACCCCGCGAGGAGGA	6580
QY	784	GACCGGATCTGTACCAACACATCTCTGTTCTTTGGGACCCCGAAGTGTACATCATC	843
Db	6581	GACCCYATTTCTATACCAACACCTATYCTGATTTTTCGGTCAACCTGTAATTTATTT	6640
QY	844	ATTCTGCGCGCTTTGGCATCATCAGCCATGTGCTGTCGACCTTCTC---GAAAGCCG	900
Db	6641	ATCTTACAGCTTCGGAATAATCTCCATATTTGTAACYTACTACTCCGGAAAAAAGAA	6700
QY	901	GTCTTCGGTTACCTGCGCATGGTCTATGCAATGGTGGCAATCGGTGTTCTGGGTTTGT	960
Db	6701	CCATTTGGATAYATAGGYATGGTCTGAGCTATATATCAATTTGGCTTCTCTGGGTTTAT	6760
QY	961	GTCTGGGGACACACATGTACACCGTTGGTATGTCGCTGACCCGACCAATCTTCTTCA	1020
Db	6761	GTGTGAGRCACCAATATTTTACAGTAGGAATAGACGTAGACACACAGCATATTTCC	6820
QY	1021	CTGGCCACCATGTTGTCGGGTGCGGACCGGCAATTAAAGATCTTCTCGTGGATCGCCAG	1080
Db	6821	TCCGACCAATAATCATCGCTATCCCGACCGGCTCAAAGTATTTAGCTGACATGCCACA	6880
QY	1081	ATGTGGGGCGCTCGGTGATGTTCAATGCCGATGCTCTGGGCTTGGCTTATGTTTC	1140
Db	6881	CTCACGGAACAAATATGAAATGATGCTGTCAGTGTCTGAGCCCTAGGATTCATTTT	6940
QY	1141	CTGTTTACCGGTGGTGTGACCGGTATCGTGTGCGCCCAAGCGGTCTGGACCGTGA	1200
Db	6941	CTTTTACCGTAGTGGCTTACTGGCATTTGTTAGCAAACTCATCTAGACATCGTA	7000
QY	1201	TATCAGGACCTATTCAGTGGTGGCGACTTCAATTATGATGTCGCTGGTGGGATC	1260
Db	7001	CTACACGACGCTACTAGTTGTAGCYCACTTCCACTATGTCCTTCAATGRRGCGWGA	7060
QY	1261	TTTGGATCTTTCGCGGTATCTACTTTTACATGCGGAAGTTCTCGGGCGCGCTTCCCG	1320
Db	7061	TTTGGCATCATAGRGCGCTTCAATCACTGATTTCCCTTATTTCTCAGGCTACACCTAG	7120

Db 861 CTTTTCGTTTCATG-----ACGTACTGATCGCGGTGCGGACCGGGATCAAGTTCTTCAA 914
Qy 1068 GTGATCGCCAGATGTGGGGCGGTGCGTTGAGTTCAAAATCGCGATGCTCTGGCCCTT 1127
Db 915 CTGGATCGCAGATGTGGAAGGGCAGTTGACATTTGAGACGGCGATGCTGTTCTGGGT 974
Qy 1128 TGGCTTTATGTTCTGTTTACCGTGGTGGTGTGACCGGTATCGTCTGGCCCAAGCGGG 1187
Db 975 CGGCTTCCTGTCACCTTCTGCTGGTGGTGTGACCGCGGTGATGCTGGCCAGCCGCC 1034
Qy 1188 TGTGACCGTGTATATCAGACACCTATTACGTGTGGCGACCTTCCCATTTATGATGTC 1247
Db 1035 GTGGACCTTCCAGTCACCGACACCTATTTCGTGGTGGCGACCTTCCACTACGTGCTGT 1094
Qy 1248 GTGGGTGGATCTTTGGATCTTGGCGGTATCTTACATGCGGAAGTTCTCGGG 1307
Db 1095 CGGCACCAATCGTGTTCGCCACCTTCGCGGGGTCTACTTCTGGTTCCCGAAGATACCGG 1154
Qy 1308 CGCGCTTTCGCGAATGGGCTGCAAGGTGCACTTCTGGACCTTCTTCATCGGTGCGAA 1367
Db 1155 CGGCTGCTGACGAGCGGCTGGGCAAGTGCATTTCTGTTGACGTTTCATCGGTTCCA 1214
Qy 1368 GTGACGTTTCTCCCGACACCTTCTGCGACGTGAGGATGCGCGCGGTTCATCGA 1427
Db 1215 CACCACCTTCTGTCAGCACTGCTGGCGGATTTGGGCATGCGCGCGCTACGCGGA 1274
Qy 1428 CTATC-----CGAAGCCTTCGCGTGTGGAAAGTCTGCTCTATGTTGCGTTCCT 1481
Db 1275 CTACCTGCCAGGAGCGGTTTCAGCCCTTACAGCTGCTCCACGCTGCGGGCGTTTAT 1334
Qy 1482 GGCCTTCGCTCTGTTCTCTTCTCATCGTATCTTTGTCTATACGCTGTTGTGTCGCG 1541
Db 1335 CTTGGCGGCTCGATGTTCCCGTTCGTCTGGAACGTCTTCAAGAGTGGCGCTACGGCGA 1394
Qy 1542 CGCGAGACCGTCCGAACCGTGGGCGAATTCGCGATAGCTGGAATGACCTGCC 1601
Db 1395 GTGCTGTCAGCTGACGACCGGTGGG-----TTACGGCAACTCGCTGGAGTGGCGACCA 1451
Qy 1602 ATCACCGCTTCGCGGCCACACGCTTCGAAACGCTGCCC 1638
Db 1452 CTCCCGCGCGCGCGGACACACTTCACCGAGCTGCC 1488

RESULT 8
AAS75523
ID AAS75523 standard; cDNA; 1555 BP.
AC AAS75523;
XX
DT 13-FEB-2002 (first entry)
DE
DE DNA encoding novel human diagnostic protein #11327.
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
OS
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Dmanac RT, Liu C, Tang YT;
PI
XX WPI; 2001-639362/73.
DR P-PSDB; ABG11336.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
P8 Claim 1; SEQ ID NO 11327; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activities. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1555 BP; 416 A; 469 C; 253 G; 417 T; 0 U; 0 Other;
Query Match 24.2%; Score 405; DB 5; Length 1555;
Best Local Similarity 57.9%; Pred. No. 4e-98;
Matches 740; Conservative 0; Mismatches 535; Indels 4; Gaps 1;
Qy 244 AACGCAACCTGTGGACGTCATGTTACCTACCATGTTATCTGATGATGTTCTTTGTG 303
Db 151 AACGACCACATCTACAACAGTTATGTCACAGCCCATGTCATTGTAATAATCTTCTTATA 210
Qy 304 GGTATCCCGCATTTGTTGGTGGTTTGGTAACATATCTGATGCGCTGCAAAATCGGCGCT 363
Db 211 GTAATACCATCATTAATCGGAGGCTTTGGCACTAGTTCCTCCCTAATAATCGGTGCC 270
Qy 364 CCGGATATGCGCTTCCGCGGTATGAACAACTGTCGTTCTGGTGTTCATTCGCGGTACC 423
Db 271 CCGGATATGCGCTTCCGCGCATAAACAATAAGCTTCTGACTCTTACCTCCCTCTCTC 330
Qy 424 CGGATGGCGTGGCTTCGCTGTTCCGACCGGGGTGACGCTGAGTGGGTTCGGGCGTT 483
Db 331 CTACTCTGCTCGCATGCTGCTATAGTGAAGCGCGGAGCAGGAACAAGTTTGAACAAGT 390
Qy 484 GGTGGGTCTGTATACCGCGCTGTGACCCGCGAAGCTGGCTATTTCGATGGACCTCGCG 543
Db 391 CGTACCTCCCTTTAGCAGGAACTACTCCACCTTTGGAAGCTCCGCTAGACCTAAC 450
Qy 544 ATTTTCGGGTTCATCTGTCGGGTGCTCTCGCATATGCGCGCGATCAATCATGATCAG 603
Db 451 ATCTCTCTTACACCTAGCAGGTGCTCTCTCTATCTTAGGGGCCATCAATTCATCACA 510
Qy 604 ACCTTCTTGAACATGCGCGCCCGCATGACGCTGCACAAAGTCCGCTGTTCTTCGTGG 663
Db 511 ACAATATCAATATAAAACCCCTGCGCATAAACCAATACCAACGCCCTCTTCGTCTGA 570
Qy 664 TCGATCTTTATCAGCGCTTGGCTGATCTGCTGGCGTTCGGGTCTGCGTGGTGGATC 723
Db 571 TCCGTCTCTAATCACAGCAGTCTCTACTTCTCTCTATCTCTCCAGTCCCTAGTGTGGCATC 630
Qy 724 ACCATGCTGTCGACCGCGTAACCTTCGGCACGACCTTCTTCAATCTGCTGGCGCGGT 783
Db 631 ACTATACTACTAACAGACCGCAACCTCAACACACCTTCTTCGACCCCGCGGAGGAGA 690
Qy 784 GACCCGATTTCTGTACCAACACATCTCTGTGGTCTTTTGGGACACCCGGAAGTGTATCATC 843

Db 691 GACCCCATTTATACCAACCTATTCTGATTTTTCGGTCACCCCTGAAGTTTATATCTT 750
 Qy 844 ATTCCTGCCGGCTTGGCATCATCAGCCATGTCTGTTTTCGACCTTTCGAAAAAGCC 899
 Db 751 ATCTACACGGCTTCGAAATATCTCCATATGTTAACTTACTCTCCGAAAAAAGA 810
 Qy 900 GGTCTTCGGTTACTCGCGATGCTGATGCAATGGTGGCAATCGTGTTCGGCTTTGT 959
 Db 811 ACCATTTGGATACATAGTATGCTGCTGAGCTATGATCAATTCGGTTCTAGGGTTAT 870
 Qy 960 CGTCTGGCGCACCACATGATACCGCTGCTGATGCTGCTGACCCAGCAATCTACTTCAT 1019
 Db 871 CGTGTGAGCACCATATATTTACAGTAGAATAGAGCTAGACACAGCATATTTTAC 930
 Qy 1020 GCTGGCCACCATGATGTCGGGTGCGGACCGCATTAAGATCTTCTGCTGGATCGCCAC 1079
 Db 931 CTCGGCTACCATATATCATGCTATCTCCCGCGCTCAAAAGTATTTAGCTGACTGCCAC 990
 Qy 1080 GATGTGGGGCGCTCGTTGAGTTCAATCGCGATGCTCTGGCGCTTTGGCTTTATGTT 1139
 Db 991 ACTCCACGGAAGCAATATGAAATGATCTGCTGAGTCTCTGAGCCCTAGGATTCATCT 1050
 Qy 1140 CTTGTTACCGTGGGTGCTGACCGGTATCGTCTGCGCCCAAGCGGGTCTGACCGTGC 1199
 Db 1051 TCTTTTACCGTAGTGGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1110
 Qy 1200 ATATCAGACACCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1259
 Db 1111 ACTACACGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1170
 Qy 1260 CTTTGGCATCTCGCGGTATCTACTTTTACATGCGGAAGTCTCGCGCGCGCTTTCC 1319
 Db 1171 ATTTGCTCATAGAGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1230
 Qy 1320 GGAATGGGTGAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1379
 Db 1231 CCAAACTACGCAAAATCCATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1290
 Qy 1380 CCGCAGCAGCTTCTGCGAGCTGAGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1439
 Db 1291 CCGACACATTTCTCGGCTATCCGGAATGCGGACGTTTCTCGGACTACCCCGATGC 1350
 Qy 1440 CTTGCGCTGTGGAACAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1499
 Db 1351 ATACACCATGAATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1410
 Qy 1500 GTTCTTCATCGTATCTTT 1518
 Db 1411 ATTAATAATTTTCATGATT 1429

RESULT 9
 ADD42976
 ID ADD42976 standard; DNA; 16559 BP.
 AC ADD42976;
 XX
 DT 15-JAN-2004 (first entry)
 XX Human mitochondrial DNA (mtDNA) SEQ ID NO:150.
 DE
 XX
 KW mitochondrial haplogroup; mitochondrial DNA; mtDNA;
 KW single nucleotide polymorphism; SNP; genetic relationship; antidiabetic;
 KW nontropic; neuroprotective; cytoskeletal; gene therapy; genealogy;
 KW forensic; Alzheimer's disease; cancer; type 2 diabetes mellitus; human;
 KW ds.
 OS Homo sapiens.
 XX
 XX WO2003046225-A1.
 XX
 PD 05-JUN-2003.

XX 25-NOV-2002; 2002MO-US038276.
 XX 26-NOV-2001; 2001US-0333622P.
 PR 28-MAR-2002; 2002US-0369131P.
 PR 01-APR-2002; 2002US-0369539P.
 XX (MITO-) MITOKOR.
 XX Herrnstadt C;
 PI WPI; 2003-505214/47.
 DR
 XX
 XX Determining single nucleotide polymorphisms in mtDNA or homoplasmic mtDNA
 PT mutations, useful for diagnosing and treating diseases, such as
 PT Alzheimer's disease, cancer and type 2 diabetes mellitus.
 XX Example 2; SEQ ID NO 150; 193pp; English.
 XX The present invention describes a method (M1) for determining the
 CC mitochondrial haplogroup of a subject, comprising determining in a
 CC biological sample with mitochondrial DNA (mtDNA) from a subject, the
 CC presence or absence of at least one mitochondrial single nucleotide
 CC polymorphism (SNP) that is associated with a mitochondrial haplogroup.
 CC Also described: (1) determining a genetic relationship between two
 CC subjects; (2) determining a genetic relationship between an unknown
 CC source or biological subject from which an unidentified sample is
 CC obtained, and a known source or biological subject from an identified
 CC sample is obtained; and (3) determining the presence of or the risk of
 CC having a disease associated with a mtDNA SNP. Mitochondrial DNA can have
 CC antidiabetic, nontropic, neuroprotective and cytoskeletal activities, and
 CC can be used in gene therapy. M1 and compositions of the present invention
 CC are useful for detecting the presence or risk of diseases, treating such
 CC diseases, determining the haplogroup of an individual, and establishing
 CC genetic relationships between individuals for genealogical and forensic
 CC purposes. The diseases include Alzheimer's disease, cancer and type 2
 CC diabetes mellitus. The present sequence represents a full length human
 CC mtDNA sequence, which is used in the exemplification of the present
 CC invention.
 XX Sequence 16559 BP; 5117 A; 5166 C; 2177 G; 4098 T; 0 U; 1 Other;
 SQ
 Query Match 23.8%; Score 398.8; DB 10; Length 16559;
 Best Local Similarity 58.2%; Pred. No. 4.6e-96;
 Matches 744; Conservative 0; Mismatches 522; Indels 12; Gaps 2;
 Qy 244 AACGACACCTGTGGAAGCTCATGTTTACCTTACCATGTTTCTGATGATGTTCTTTGTG 303
 Db 6050 AACGACACATCTACAGCTTATGTCACAGCCCATGATTTGTAATATCTTTCATA 6109
 Qy 304 GGTATCCCGCGATTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 363
 Db 6110 GTAATACCATCATAATCGGAGGCTTTGGCACTGACTAGTTCCTTAATCGGTGCC 6169
 Qy 364 CGGATATGGCTTCCCGCTGATGAACCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 423
 Db 6170 CCGATATGGGCTTCCCGCGCATAAACATAAGTCTTCTGACTTCTACCTCCCTCTCTC 6229
 Qy 424 CGGATGGGCGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 483
 Db 6230 CTACTCTGCTCGCATCTGCTATAGTAGGCGCGGAGCAGGAGGTTGAACAGTCT 6286
 Qy 484 GTTGGGTTCTGTACCGCGCTGCTGACCGCGAAGCTGGTATTCGATGGACCTCGCG 543
 Db 6287 -----TACCTCTCTAGCAGGAACTACTCCACCTCGAGCTCCGTAGACCTAAC 6340
 Qy 544 ATTTTCGGGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 603
 Db 6341 ATCTTCTCTTACACCTAGCAGGTTGCTCTCTCTATCTTAGGGGCCATCAATTCATCACA 6400
 Qy 604 ACCTTCTTGAACATGCG 663
 Db 6401 ACAATTATCAATATAAAACCCCTTGCATACCAATACCAACAGCCCTCTTCTGCTGTA 6460

Db 6289 -----TACCTCCCTTAGCAGGAACTACTCCACCGTGGAGCCTCCGTAGACCTAAC 6342
Qy 544 ATTTTCGGGGTTCACTTGTGGGTGCTCTTCGATCATGGCGGATCAACATCATCAAG 603
Db 6343 ATCTTCTCTTACACCTAGAGGTGTCTCTCTATCTTAGGGCCATCAACTCATCACA 6402
Qy 604 ACCTTCTTGAACATCGCGGCCCGGATGACGCTGCACAAAGTCCGTTTCTCTGTTG 663
Db 6403 ACAATTATCAATATAAAACCCCTGCCATAACCAACCAACGCGCCCTCTTCTGTGA 6462
Qy 664 TCGATCTTTATCAGCGCTTGGCTGATCTGTGGCGGTGCGGTTCTGGCTGGTGAATC 723
Db 6463 TCGTCTCTAATCAGCAGAGTCTCTCTCTCTCTCTCCAGTCTAGCTGTGGCATC 6522
Qy 724 ACCATGCTGTGACCGACCGTAACTTGGCAGCAGCTTCTTCAATCTGCTGGCGCGGT 783
Db 6523 ACTATACTACTAACAGACCGCACTCAACACACCTTCTTCGACCCGCGGAGGAGGA 6582
Qy 784 GACCGGATCTGTACACACATCTCTGTGTTCTTTGGGACCGGAAAGTGTATCATC 843
Db 6583 GACCCCATCTTATACCAACCTATCTGATTTTTCGGTCACTTGAAGTTTATATCTT 6642
Qy 844 ATCTGCGCGCTTGTGGCATCATCAGCCATGTCTGTGCGACCTTCTC--GAAAAAGCCG 900
Db 6643 ATCTGACCGCTTGGGAATAATCTCCATATTTGTAATCTACTCTCGGAAAGAGAA 6702
Qy 901 GTCTTCGTTACCTGCGGATGCTATGCAATGTGGCAATCGGTCTTGGGCTTTGTC 960
Db 6703 CCATTTGGATACATAGTATGCTGAGCTATGATATCAATTTGGCTTCTTAGGTTTATC 6762
Qy 961 GTCTGGCGCACCATGTACACCGTTGGTATGTGCTGACCCAGCAATCTACTTATG 1020
Db 6763 GTGTGACACACATATATTACAGTAGAATAGACGTAGACACAGCATATTTTACC 6822
Qy 1021 CTGGCCACCATGTTGATCGGGGTGCCACCGGATTAAGATCTTCTGTGGATCGCCACG 1080
Db 6823 TCCGCTACCATATCATCGCTATCCCGCGGGTCAAGTATTTAGCTGACTCGCCACA 6882
Qy 1081 ATGTGGCGCGCTGCTGTAGTTCAATCGCGATGCTGGGCTTTGGCTTTATGTTTC 1140
Db 6883 CTCACGAGCAATATGAATGATCTGTGCGAGTCTCTGAGCCCTTAGGATTCATCTT 6942
Qy 1141 CTGTTCACCGTGGTGTGACCGGTATCGTGTGCGCCCAAGCGGTCTGACCGTGA 1200
Db 6943 CTTTTCACCGTAGTGGCTGACTGGCATTTGATTAGCAAACTCATCACTAGACATCGTA 7002
Qy 1201 TATCAGCACACCTATTACGTGGTGGGCATCTTCATATGTGATGTGCTGGGTGGATC 1260
Db 7003 CTACACGACACGTACTAGTTGAGTCACTTCCATCTATCTCATTAAGAGCTGA 7062
Qy 1261 TTTTGCATCTTTCGCGGTATCTACTTTTACATCGCGAAGTTCTCGGCGCGCTTTCCCG 1320
Db 7063 TTTGCCATCATAGAGCTTCATCTACTGATTTCCCTATTCTCAGCTACACCTTAGAC 7122
Qy 1321 GAATGGCTGCAAGCTGCACTTGTGAACCTTCTCATCGTGGCAAGTCACTTCTTC 1380
Db 7123 CAACCTTACGCCAAAATFCAATTTCACTATCATATTCATCGGCGTAAATCTAATCTTTC 7182
Qy 1381 CCGCAGCATTCTCGGGAGCTCAGGATGATCGCGCGGTATACATCGACTATCCGAGCC 1440
Db 7183 CCACACATCTTCTCGCCTATCCGNAATGCCCGAGCTTACTCGGACTACCCCGATGA 7242
Qy 1441 TTGCGCTGTGGAAACAAAGTCTCTGTCTATGTGCGTCTCTGCGCTTCCGCTGTTCTG 1500
Db 7243 TACACCACATGAACATCTCTATCATCTGTAGGCTCATCTTCTCTAACAGCAGTAATA 7302
Qy 1501 TTCTTCATCGTATCTTT 1518
Db 7303 TTAATAATTTTCATGATT 7320

ADD43240
ID ADD43240 standard; DNA; 16568 BP.
XX AC ADD43240;
XX DT 15-JAN-2004 (first entry)
XX Human mitochondrial DNA (mtDNA) SEQ ID NO:414.
XX mitochondrial haplogroup; mitochondrial DNA; mtDNA;
KW single nucleotide polymorphism; SNP; genetic relationship; antidiabetic;
KW neoplastic; neuroprotective; cytosolic; gene therapy; genealogy;
KW forensic; Alzheimer's disease; cancer; type 2 diabetes mellitus; human;
ds.
XX Homo sapiens.
OS
XX WO2003046225-A1.
PN
XX 05-JUN-2003.
PD
XX 25-NOV-2002; 2002WO-US038276.
PF
XX 26-NOV-2001; 2001US-0333622P.
PR
XX 28-MAR-2002; 2002US-036913P.
PR
XX 01-APR-2002; 2002US-0369539P.
XX (MITO-) MITOKOR.
XX Herrnstadt C;
XX WPI; 2003-505214/47.
XX
XX Determining single nucleotide polymorphisms in mtDNA or homoplasmic mtDNA
mutations, useful for diagnosing and treating diseases, such as
PT Alzheimer's disease, cancer and type 2 diabetes mellitus.
XX
XX Example 2; SEQ ID NO 414; 193pp; English.
XX
XX The present invention describes a method (M1) for determining the
CC mitochondrial haplogroup of a subject, comprising determining in a
CC biological sample with mitochondrial DNA (mtDNA) from a subject, the
CC presence or absence of at least one mitochondrial single nucleotide
CC polymorphism (SNP) that is associated with a mitochondrial haplogroup.
CC Also described: (1) determining a genetic relationship between two
CC subjects; (2) determining a genetic relationship between an unknown
CC source or biological subject from which an unidentified sample is
CC obtained, and a known source or biological subject from an identified
CC sample is obtained; and (3) determining the presence of or the risk of
CC having a disease associated with a mtDNA SNP. Mitochondrial DNA can have
CC antidiabetic, neoplastic, neuroprotective and cytosolic activities, and
CC can be used in gene therapy. M1 and compositions of the present invention
CC are useful for detecting the presence or risk of diseases, treating such
CC diseases, determining the haplogroup of an individual, and establishing
CC genetic relationships between individuals for genealogical and forensic
CC purposes. The diseases include Alzheimer's disease, cancer and type 2
CC diabetes mellitus. The present sequence represents a full length human
CC mtDNA sequence, which is used in the exemplification of the present
CC invention.
XX
SQ Sequence 16568 BP; 5116 A; 5186 C; 2176 G; 4089 T; 0 U; 1 Other;

Query Match 23.8%; Score 398.8; DB 10; Length 16568;
Best Local Similarity 58.2%; Pred. No. 4.6e-96;
Matches 744; Conservative 0; Mismatches 522; Indels 12; Gaps 2;
Qy 244 AACGACACCTGTGGAACGTCATGTTACCTACATCGTATCTGATGATGTTCTTGTG 303
Db 6050 AACGACACATCTTCAACGTTATCGTCACAGCCATGATTTGTAATAATCTTCTTCATA 6109
Qy 304 GGATATCCCGCATTTGTCGGTGGTTTGGTAACTATCTGATCGCGTGAATCGCGCT 363
Db 6110 GTAATACCATCATATAATCGGAGGCTTTGGCACTGACTAATCCCTAATAATCGGTGCC 6169

QY	364	CGGATATAGCCTTCCGGGTATGAACAACTGTGTTCTGGTGTTTCATTTGCCGGTACC	423
Db	6170	CCGGATATGGCGTTTCCCGCATAAACAAATAGCTTCTTGACTCTTACCTCCCTCTCTC	6229
QY	424	GCGATGGCGTGGCTTCCGCTGTTCCGACCCGGGGGTGACGGTCAGCTGGGTTCGGGGCTT	483
Db	6230	CTACTCTGCTCGCATCTGCTATAGTGGAGCCGGAGCAGAAACAGGTTGAAACAGTC---	6286
QY	484	GGTTGGGTTTGTATACCCGCGCTGTGCACCGCGAAGCTGGCTATTTCGATGGACCTCGCG	543
Db	6287	-----TACCCTCCCTTAGCAGGAACTACTCCACCCCTGGAGCTCCGAGACCTAAACC	6340
QY	544	ATTTCGGGTTACATTTGTCCGGTGCCCTCTCGATCATGGCGCGATCAACATGATCAAG	603
Db	6341	ATCTTCTCCTTACACCTAGCAGGTGTCTCTCTATCTTAGGGCCATCAATTTCAATCACA	6400
QY	604	ACCTTCTTTGAACATGCGCGCCCGGATGACGCTGCACAAAGTGGCGTTGTCTCGTGG	663
Db	6401	ACAAATATCAATATAAAACCCCTGCCATAACCAATACCAACAGCCCTCTTCGTCTGA	6460
QY	664	TCGATCTTTATACGGCTTGGCTGATCTCTCTGGCGCTGCCGTTTGGCTGTGGAATC	723
Db	6461	TCCGTCCTAATACAGCAGTCTTACTTCTCTATCTCTCCAGCTCTAGCTGTGGCATC	6520
QY	724	ACCATGTGCTGACCGACCGTAACTTCGGCACACCTTCTTCAATCTGCTGGCGCGGT	783
Db	6521	ACTATACTACAGACCGCACTCAACACACCTTCTTCGACCCCGCGGAGAGGA	6580
QY	784	GACCGAATCTGTACCAACACATCTGTGTTCTTTGGGACCCGGAGTGATCATCATC	843
Db	6581	GACCCCAATCTATACCAACACCTATTCTGATTTTTCGGTCACCCCTGAAGTTTATATCT	6640
QY	844	ATTCTGCCGGCTTTGGCATCATCAGCAATGTCTGTCCGACCTTCTC--GAAAAGACCG	900
Db	6641	ATCTACACGGCTTCGGAAATACTCCCATATTGTAACTTTACTTACTCCGGAATAAGAA	6700
QY	901	GTCTTCGGTTACCTGCCGATGCTATGCAATGGTGGCAATCGGTTCGTGGCTTTGTC	960
Db	6701	CCATTTGGATACATAGTATGGTCTGAGCTATGATATCAATGGCTCTTAGGGTTTATC	6760
QY	961	GTCTGGCGCACCATGTATACCGTTGGTATGTGCTGACCCAGCAATCCTACTTCATG	1020
Db	6761	GTGTGAGCACCATATATTACAGTAGGAATAGAGCTAGACACAGAGCATATTTCAACC	6820
QY	1021	CTGSCCAACCAATGGTATCGCGGTGCCAGCGGATTAAGATCTTCTGTGGATCGGCACG	1080
Db	6821	TCGCTACCATATCATCGTATCCCCACCGGGTCAAGATATTAGCTGACTCGCCACA	6880
QY	1081	ATGTGGGGCGGCTGGTTGAGTTCAAAATCGCGAAGCTCTGGGCTTTTGGCTTTATGTTT	1140
Db	6881	CTCACGGAAAGCAATATGAATGATCTGTGCAAGTGTCTGAGCCCTAGGATTCATCTTT	6940
QY	1141	CTGTTCACCGTGGTGTGACCGGTATCGTGTGCCCAAGCGGTCTGGAACCGTGA	1200
Db	6941	CTTTTCAACGTAGTGGCTTACTGGCATTTGATTAGCAACTCATCATAGACATCGTA	7000
QY	1201	TATCAGACACCTATTACGTGGTGGCGCACTTCGATTGTGATGTGCTGGGTGGGATC	1260
Db	7001	CTCACAGACAGTACTAGTTTGAAGTCACTTCCACTATGTCTCTATCAATAGGAGCTGA	7060
QY	1261	TTTTCGATCTTCGCGGTATCTACTTTTACATGCGAAAGTTCTCGGCGCGGCTTTCCCG	1320
Db	7061	TTTGGCATCATAGAGGCTTCATTCACTGATTTCCCTATTCTCAGGCTACACCTAGAC	7120
QY	1321	GAATGGGCTCGAAGCTGCATCTTGACCTTCTTCATCGTGGAGACGTCAAGTTCTTTC	1380
Db	7121	CAAACTTACGCAAAATCCATTTCACTATCATATTTATCGGCGTAAATCTAATTTCTTC	7180
QY	1381	CCGACGACCTTCTGGGACGTCAAGGTATGCCGCGCGTTTACATCCGACTATCCGAAGC	1440
Db	7181	CCACAAACATTTCTCGGCTTATCCGAAATGCCCGGAGCTTACTCGGACTACCCCGATGA	7240

Qy	1441	TTCCGGCGTGGGAACAAAGTCTCGTCTATGCTATGGTGGCTTCCTGGCCCTTCGCTCGTTCCTG	1501
Db	7241	TACACCACATGAACATCCCTATCATCTGTAGGCTCATTCATTTCTCTAACAGCAGTAATA	7300
Qy	1501	TTCTTCATCGTGATCTTT	1518
Db	7301	TTAATAATTTTCATGATT	7318
RESULT 12			
ADD43266	ID	ADD43266 standard; DNA; 16569 BP.	
XX	AC	ADD43266;	
XX	DT	15-JAN-2004 (first entry)	
XX	DE	Human mitochondrial DNA (mtDNA) SEQ ID NO:440.	
XX	KW	mitochondrial haplogroup; mitochondrial DNA; mtDNA;	
KW	KW	single nucleotide polymorphism; SNP; genetic relationship; antidiabetic;	
KW	KW	neurotropic; neuroprotective; cytotstatic; gene therapy; genealogy;	
KW	KW	forensic; Alzheimer's disease; cancer; type 2 diabetes mellitus; human;	
XX	XX	ds.	
XX	OS	Homo sapiens.	
XX	PN	WO2003046225-A1.	
XX	PD	05-JUN-2003.	
XX	PF	25-NOV-2002; 2002WO-US038276.	
XX	PR	26-NOV-2001; 2001US-033622P.	
PR	PR	28-MAR-2002; 2002US-036931P.	
PR	PR	01-APR-2002; 2002US-0369539P.	
XX	PA	(MITO-) MITOKOR.	
XX	PI	Hexrnstadt C;	
XX	PI	WPI; 2003-505214/47.	
XX	DR	Determining single nucleotide polymorphisms in mtDNA or homoplasmic mtDNA	
PT	PT	mutations, useful for diagnosing and treating diseases, such as	
PT	PT	Alzheimer's disease, cancer and type 2 diabetes mellitus.	
XX	PS	Example 2; SEQ ID NO 440; 193pp; English.	

The present invention describes a method (M1) for determining the mitochondrial haplogroup of a subject, comprising determining in a biological sample with mitochondrial DNA (mtDNA) from a subject, the presence or absence of at least one mitochondrial single nucleotide polymorphism (SNP) that is associated with a mitochondrial haplogroup. Also described: (1) determining a genetic relationship between two subjects; (2) determining a genetic relationship between an unknown source or biological subject from which an unidentified sample is obtained, and a known source or biological subject from an identified sample is obtained; and (3) determining the presence of or the risk of having a disease associated with a mtDNA SNP. Mitochondrial DNA can have an anti-diabetic, neurotropic, neuroprotective and cytosstatic activities, and can be used in gene therapy. MI and compositions of the present invention are useful for detecting the presence or risk of diseases, treating such diseases, determining the haplogroup of an individual, and establishing genetic relationships between individuals for genealogical and forensic purposes. The diseases include Alzheimer's disease, cancer and type 2 diabetes mellitus. The present sequence represents a full length human mtDNA sequence, which is used in the exemplification of the present invention.

Sequence 16569 BP; 5115 A; 5174 C; 2177 G; 4102 T; 0 U; 1 Other;
Query Match 23.8%; Score 398.8; DB 10; Length 16569;

Best Local Similarity 58.2%; Pred. No. 4.6e-96;
Matches 744; Conservative 0; Mismatches 522; Indels 12; Gaps 2;

```
QY 244 AACGACACCTGTGGAAGTGTATGTTACCTACATGATGTTCTGATGATGTTCTTTGTG 303
Db 6051 AACGACACACATCTCAACAGTTATGTCACAGCCATGATTTGTAATAATCTTTTCATA 6110
QY 304 GGTATCCCGCATGTTTCGGTGGTTTGGTAACTATCTGATGCCGTCGCAATCGGCGCT 363
Db 6111 GTATACCATCATAAATCGGAGGCTTTGGCACTGATGTTCCCTTAATAATCGGTGCC 6170
QY 364 CGGATATGGCTTCCCGCATGATGAAACAACTGCTGTTCTGGCTGTTTCATITCGCGTACC 423
Db 6171 CCGATATGGGTTTCCCGCATAAACAACTAAAGCTTCTGACTCTTACCTCCCTCTCTC 6230
QY 424 CGGATGGCGGTGCTTGGCTGTCGACCGGCGGTGACGCTGAGCTGGGTTGGGCGTT 483
Db 6231 CTACTCTGCTCGATCTGCTATGATGAGCGCGGAGCGGAGCAAGGTTGAAAGTCTC 6287
QY 484 GGTGGGTTCTGTACCGCGCTGTGACCGCGAAGCTGGCTATTCGATGGACCTCGCG 543
Db 6288 -----TACCCCTCTTAGCAGGAACTACTCCACCCCTGGAGCTTCGTAGACCTAAC 6341
QY 544 ATTTTCGGGTTCACTTGTGGGTGCTCTCTCGATCATGGGCGGATCAACATGATCAAG 603
Db 6342 ATCTCTCTTACACCTAGCAGGTTCTCTCTATCTTAGGGGCCATCAATTTTCATCACA 6401
QY 604 ACCTTCTTGAACATGCGCGGCCCCCGGATGACGCTGCAAAAGTGCCTGTTGTTCTGTTG 663
Db 6402 ACAATTATCAATAAARACCCCTGCAATAACCAATACCAAGCCCTCTTCTGCTGA 6461
QY 664 TCGATCTTATCAGCGTTGGCTGATCCCTGCTGGCGTGCCTGTTGCGGTGCTGCAATC 723
Db 6462 TCCGTCCTTAATACAGCAGTCTACTTCTCTATCTCTCCAGTCTAGTGTGGCATC 6521
QY 724 ACCATGCTGTGACCGACCGTAACTTCGGCACGACCTTCTCAATCTGCTGCGCGGCT 783
Db 6522 ACTATATCTACAGACCGCACTCAACACACCTTCTTCACCCCGCGGAGGAGA 6581
QY 784 GACCCGATCTGTACCAACACATCTGTGTTCTTTGGGCAACCGGAAGTGTACATCATC 843
Db 6582 GACCCCATCTATACCAACACCTATCTGATTTTTCGCTCACCTCAAGTTTATATCTT 6641
QY 844 ATCTGCGCGCTTTGSCATCATCAGCATGCTGCTGCTGACCTTCTC---GAAAAGCG 900
Db 6642 ATCTACACGCTTCGAATAATCTCCATTTGTAACTTACTTACCCGGAAGAAAGAA 6701
QY 901 GTCTTGGTTTACCTGCGGATGTTCTATGCAATGTTGGCAATCGGTTCTTGGGCTTTGTC 960
Db 6702 CCATTTGGATACATAGTATGTTGCTGAGTATGATATCAATTTGGCTTCTAGGGTTATC 6761
QY 961 GTCTGGGCGCACCATGTACACCGTTGATGTCGCTGACCCAGCAATCTTACTTATG 1020
Db 6762 GTGTGAGCACCATATATTTACAGTAGAATAGACGTAGACACAGCATATTTTACC 6821
QY 1021 CTGCGCACCATGTTGATCGCGTCCCGGCAATTAAGATCTTCTCGGTGATTCGCCAG 1080
Db 6822 TCCGCTACCATATATCATCGTATCCCAACCGGCTCAAGTATTTAGTGTACTGCCACA 6881
QY 1081 ATGTGGGCGGCTCGGTTGAGTTCAATCGCGATGCTCTGGGCTTTGGCTTTATGTT 1140
Db 6882 CTCACGGAAGCAATATGAATGATGCTGTCGAGTCTCTGAGCCCTAGGATTTATCTT 6941
QY 1141 CTGTTACCGTGGTGTGACCGGTATCGGTGCTGGCCCAAGCGGTTCTGACCGTGA 1200
Db 6942 CTTTTCACCGTAGGTGCTGACTGTCATGTTATGACCAATCATCATAGACATGTA 7001
QY 1201 TATCAGCACCTATTTACGTGGGCGCACTTCCATTTATGATGTCGTGGGTGCGATC 1260
Db 7002 CTACACGACACGCTACTACGTTGTAGTCTACTTCCACTATGCTCTATCAATAGGAC 7061
QY 1261 TTTGGCATCTTCCGCGTATCTACTTTTACATGCGGAGTCTCGGCGGCGCTTTCCG 1320
```

```
Db 7062 TTTGCCATCATAGGAGGTTTCATTCACTGATTTCCCTTATTTCTCAGGCTACACCTAGAC 7121
QY 1321 GAATGGGCTGCAAAAGTGCACCTTCTGGACCTTCTTCATCGGTGCGAAGTCACTGTTCTTC 1380
Db 7122 CAAACCTACGCCAAATCCATTCTACTATCATTAATTCATCGGGTAAATCTAACTTCTTC 7181
QY 1381 CGCGACACTTCTCGGAGCTGAGGATGCGCGGCTTACATGATATCCCGAAGCC 1440
Db 7182 CCACAACACTTCTCGGCTATCCGGAATGCCCGACGTTACTCGGACTACCCGATGCA 7241
QY 1441 TTTGGGCTGTGAAACAAAGTCTCTGCTATGTTGGTTCTCGGCTTCGCTCGTTCTG 1500
Db 7242 TACACCACATGAAACATCTATCATCTGTAGGCTCATTCATTTCTCTAACAGCATATA 7301
QY 1501 TTCTTCATCGTGATCTTT 1518
Db 7302 TTAATAATTTTCATGATT 7319
```

RESULT 13

ADP43327

ID ADD43327 standard; DNA; 16569 BP.

XX AC ADD43327;

XX AC ADD43327;

XX DT 15-JAN-2004 (first entry)

XX Human mitochondrial DNA (mtDNA) SEQ ID NO:501.

DE Human mitochondrial DNA (mtDNA) SEQ ID NO:501.

XX single nucleotide polymorphism; SNP; genetic relationship; antidiabetic;

KW neotrophic; neuroprotective; cytostatic; gene therapy; genealogy;

KW forensic; Alzheimer's disease; cancer; type 2 diabetes mellitus; human;

XX ds.

XX Homo sapiens.

XX WO2003046225-A1.

XX C5-JUN-2003.

XX 25-NOV-2002; 2002WO-US038276.

XX 26-NOV-2001; 2001US-0333622P.

XX 28-MAR-2002; 2002US-0369131P.

XX 01-APR-2002; 2002US-0369539P.

XX (MITO-) MITOKOR.

XX Hernstadt C;

XX WPI; 2003-505214/47.

XX Determining single nucleotide polymorphisms in mtDNA or homoplasmic mtDNA

XX mutations, useful for diagnosing and treating diseases, such as

XX Alzheimer's disease, cancer and type 2 diabetes mellitus.

XX Example 2; SEQ ID NO 501; 193pp; English.

XX The present invention describes a method (M1) for determining the

XX mitochondrial haplogroup of a subject, comprising determining in a

XX biological sample with mitochondrial DNA (mtDNA) from a subject, the

XX presence or absence of at least one mitochondrial single nucleotide

XX polymorphism (SNP) that is associated with a mitochondrial haplogroup.

XX Also described: (1) determining a genetic relationship between two

XX subjects; (2) determining a genetic relationship between an unknown

XX source or biological subject from which an unidentified sample is

XX obtained, and a known source or biological subject from an identified

XX sample is obtained; and (3) determining the presence of or the risk of

XX having a disease associated with a mtDNA SNP. Mitochondrial DNA can have

XX antidiabetic, neotrophic, neuroprotective and cytostatic activities, and

XX can be used in gene therapy. M1 and compositions of the present invention

XX are useful for detecting the presence or risk of diseases, treating such

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC diseases, determining the haplogroup of an individual, and establishing
CC genetic relationships between individuals for genealogical and forensic
CC purposes. The diseases include Alzheimer's disease, cancer and type 2
CC diabetes mellitus. The present sequence represents a full length human
CC mtDNA sequence, which is used in the exemplification of the present
XX invention.

SQ Sequence 16569 BP; 5118 A; 5184 C; 2175 G; 4091 T; 0 U; 1 Other;

Query Match 23.8%; Score 398.8; DB 10; Length 16569;
Best Local Similarity 58.2%; Pred. No. 4.6e-96;
Matches 744; Conservative 0; Mismatches 522; Indels 12; Gaps 2;

QY 244 AACGGACACCTGTGGAAAGCTCATGTTACCTACCAATGTTATTCGATGATGTTCTTTGTG 303
DB 6051 AACGACACATCTACACAGTTATCGTCACAGCCCATGATTTGTAATAATCTTTCATA 6110
QY 304 GGTATCCCGGCTTGTGGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 363
DB 6111 GTAATACCATCATTAATCGGAGGCTTGGCAACTGACTAGTTCCTCCCTAATTAATCGTGCC 6170
QY 364 CCGGATATGGCCTTCCCGGATATGAACAACTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 423
DB 6171 CCCGATATGGCGTTTCCCGGCAATAACACATAAGCTTCTGACTCTTACCTCCCTCTCTC 6230
QY 424 GCGATGGGCGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 483
DB 6231 CTACTCTCTCTGCT 6287
QY 484 GGTGGGTTCTGTACCGCGCGTGTCTGACCCGCAAGCTGGCTATTCGATGACCTCGCG 543
DB 6288 -----TACCTCTCTCTACAGGAACTACTCCACCTGGAGCTCGTGACCTAAC 6341
QY 544 ATTTTCGGGTTTCACTGTTCGGTGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 603
DB 6342 ATCTTCTCTTACACCTAGCAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 6401
QY 604 ACCTTCTTGAACATGCGCGCGCGCGCGCATGCGCTGCACAAAGTGGCTGTCTCTCTCTCTCT 663
DB 6402 ACAATATCAATATAAAACCCCTCGCATACCCAAATACAAAGCCCTCTCTCTCTCTCTCTCT 6461
QY 664 TCGATCTTATCACGGCTTGGTGTGATCTCTGTGGCGCTCCGCTCTCTCTCTCTCTCTCTCTCT 723
DB 6462 TCGTCTCTAATACAGAGTCT 6521
QY 724 ACATCTGTGTACCGACGTAATCTCTGCGGACGACCTTCTCAATCTCTCTCTCTCTCTCTCT 783
DB 6522 ACTATACTACTACAGACCGCAACCTCAACACCACTTCTCTGACCCCGCGGAGGAGA 6581
QY 784 GACCGATCTGTACCAACACATCTCTGTGGTCTTTGGGCAACCGGAAAGTGTACATCATC 843
DB 6582 GACCCCACTCTATACCAACACCTATCTGATTTTTCGGTCACTGAGTTTATATTTCT 6641
QY 844 ATTCTGCGCGCTTGTGATCATCATGCGCATGTCTGTGTGACCTTCTCTCTCTCTCTCTCTCT 900
DB 6642 ATCTTACCGCTTGGGAATAATCTCTCCATATTGTAACTTACTCTCCGGAATAAAGAA 6701
QY 901 GTCTTGGTTACCTGCGGATGTTCTATGCAATGTTGGCAATGTTCTGGGCTTGTCT 960
DB 6702 CCATTTGGATACATAGTATGCTGTAGCTATGATATCAATTTGCTTCTAGGTTTATC 6761
QY 961 GTCTGGGCGCACCATCTACACGTTGGTATGTCGCTGACCCAGCAATCTCTACTTCATG 1020
DB 6762 GTGTGACACACCATATATTACAGTAGGAATAGACGTAGACACACAGCATATTTTCAAC 6821
QY 1021 CTGCGCACCATGTGTGTCGGGTGCGGACCGGATTAAGATCTCTCTGTCGATCGCCACG 1080
DB 6822 TCCGCCACCATTAATCATCGCTATCCCGACCGGCTCAAGTATTTAGCTGACTCGCCACA 6881
QY 1081 ATGTGGGCGCGCTCGGTGAGTTCAATTCGCGCATGCTCTGGGCTTGTGGCTTTATGTTTC 1140
DB 6882 CTCCACGGAACAATATGAATGATCTGCTGCAGTCTCTGAGCGCTTAGGATTCATCTTT 6941

QY 1141 CTGTTACCGTGGTGGTGTGACGGTATCTGCTGGCCCAAGCGGTCTGGACCGTGCA 1200
DB 6942 CTTTTCACCGTAGGTGGCTGACTGGCAATTGATTAGCAAACTCATCATAGACATCGTA 7001
QY 1201 TATCAGACACCTATTACGTGTGGCGCACTTCCCATTTATGTGATGTCTGGTGGGATC 1260
DB 7002 CTACACGACGCTACTACGTTGTAGCCCACTTCCCACTATGTCTCTATCAATAGGAGTGTA 7061
QY 1261 TTTGGGATCTTCCCGGTATCTACTTTTACATGCCCAAGTTCTCTGGGCGCGCTTTCCCG 1320
DB 7062 TTTGGCATCATAGGAGCTTCAATCTGATTTTCCCTTATTCFAGGCTACACCTAGAC 7121
QY 1321 GAATGGGCTGCAAGTGCACCTTCTGGACCTTCTCATCGGTGCGAAGCTCAGTTCTTC 1380
DB 7122 CAAACCTACGCCAAATCAATTTCACTATCAATTCATCGGCGTAAATCTAATTTCTTC 7181
QY 1381 CCGCAGACATCTCTGGGAGCTCAGGTATGCCCGCGCTTACATCGACTATCCGAGCC 1440
DB 7182 CCAACACATTTCTCGGCTTATCCGAATGCCCGACGTTACTCGGACTACCCCGATGCA 7241
QY 1441 TTGCGGCTGTGGAAACAAAGTCTGCTCTATGTGTGGTCTCTGGCCTTGGCTCTGTTCTG 1500
DB 7242 TACACCACATGAACATCCCTATCATCTGTAGGCTCATTTCTCTAACAGCAGTAATA 7301
QY 1501 TTCTTCACTCGTATCTTT 1518
DB 7302 TTAATAATTTTTCATGATT 7319

RESULT 14

ADD43295

ID ADD43295 standard; DNA; 16569 BP.

XX ADD43295;

XX 15-JAN-2004 (first entry)

XX Human mitochondrial DNA (mtDNA) SEQ ID NO:469.

XX mitochondrial haplogroup; mitochondrial DNA; mtDNA;
XX single nucleotide polymorphism; SNP; genetic relationship; antidiabetic;
XX neotropic; neuroprotective; cytosolic; gene therapy; genealogy;
XX forensic; Alzheimer's disease; cancer; type 2 diabetes mellitus; human;
XX ds.

XX Homo sapiens.

XX WO2003046225-A1.

XX 05-JUN-2003.

XX 25-NOV-2002; 2002WO-US038276.

XX 26-NOV-2001; 2001US-0333622P.

XX 28-MAR-2002; 2002US-0369131P.

XX 01-APR-2002; 2002US-0369539P.

XX (MITO-) MITOKOR.

XX Herrnsstadt C;

XX WPI; 2003-505214/47.

XX Determining single nucleotide polymorphisms in mtDNA or homoplasmic mtDNA
XX mutations, useful for diagnosing and treating diseases, such as
XX Alzheimer's disease, cancer and type 2 diabetes mellitus.

XX Example 2; SEQ ID NO 469; 193pp; English.

XX The present invention describes a method (M1) for determining the

XX mitochondrial haplogroup of a subject, comprising determining in a

XX biological sample with mitochondrial DNA (mtDNA) from a subject, the

XX presence or absence of at least one mitochondrial single nucleotide

Search completed: October 22, 2004, 14:49:47
Job time : 1007 secs

Search completed: October 22, 2004, 14:49:47
Job time : 1007 secs

QY 678 GCCTTGGCTGATCCTGCTGGCGTGGCGGTTCTGGCTGGTGAATCACCATGCTGCTGAC 737
Db |||||
QY 379 CGCGTCTGCTGATCGCGGTGATGCGGTGCTGGCGGCTGGTGAATGCTGAT 438
Db |||||
QY 738 CGACCGTAACCTGGCGACGACCTTCTCAATCTGCTGGCGGCTGACCCGATTTCTGTA 797
Db |||||
QY 439 GGACATCCACTTCGGACCAAGCTCTTCAGCGCGCGCGCGCGGCGACCGGTGCTGTT 498
QY 798 CCAACATCCTGCTGCTTCTGGCGACCGGAGTGTACATCATATTCGCGCGCTT 857
Db |||||
QY 499 CCAGCAGTGTCTGGTCTTCGGCCACCGGAGGTGTACATCATATTCGCGCGCTT 558
QY 858 TGGCATCATCAGCCATGCTGCTGCGACCTTCTCGAAAAGCCGCTCTTCGCTTACCTGCC 917
Db |||||
QY 559 CGGTGCGGTGAGTCCGATCATCCGACCTTCGCGCGCAAGCGCTGTTGGCTACACCTC 618
QY 918 GATGCTCTATGCAATGCTGGCAATCGGTGTTCTGGGCTTGTGCTGGCGCGACACAT 977
Db |||||
QY 619 GATGCTCTAGCCACCGCCACGATCGCTTCTCTCTGCTGCTGGCGCGACACAT 678
QY 978 GTACACCGTGTGATGCTGCTGACCCAGCAATCTACTTCACTGCTGCCACCATGCTGAT 1037
Db |||||
QY 679 GTTCTGCTCGGATCCGCTGTCACCGCGAGCTGTTCTTATGACCCACCATGCTGAT 738
QY 1038 CGCGGTGCGGACCGGATTAAGATCTTCTGCTGATGCGCACGATGCGGCGGCTCGGT 1097
Db |||||
QY 739 CGCGGTGCCACCGGCTGAAGGTGTTCAACTGGGTGACCCACCATGCTGGAGGTTGCT 798
QY 1098 TGAGTTCAAATCGCGATGCTGCGGCTTGTGCTTATGCTTCTGCTTACGCTGGGCTG 1157
Db |||||
QY 799 GACCTTCGACGCGGATGCTGCTGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 858
QY 1158 TGTGACCGGTATCGTGTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 1217
Db |||||
QY 859 CTTCGCGGCTGATGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 918
QY 1218 CGTGGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1277
Db |||||
QY 919 CGTGGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 978
QY 1278 TATCTACTTTTACATGCGGAGTCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 1337
Db |||||
QY 979 GGCTTACTACTGCTGCGGAGTGGACCGGCGGCTGCGGCTGCGGCTGCGGCTGCGG 1038
QY 1338 GCATCTTGGACCTTCTTCAATCGGTGGAGCTGACGTTCTTCCGCGGCTGCTGCTGCTG 1397
Db |||||
QY 1039 GCATCTTGGATGAGCTTCTATCGGAGTAACTGCGGCTGCTTCCCGATGCTGCTGCTG 1098
QY 1398 ACCTGAGGTATGCGGCGGCTTACATGCACTATCCGAGGCTTCCGCTGCTGGAACAA 1457
Db |||||
QY 1099 CTTGCGGCTGATGCGGAGTGGACCGGCGGCTGCGGCTGCGGCTGCGGCTGCGG 1158
QY 1458 AGTCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1517
Db |||||
QY 1159 GGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1215
QY 1518 TGTCTATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1577
Db |||||
QY 1216 GGTCTATCAAGTGTATGCGGCGGAGCGGCGGCTGCGGCTGCGGCTGCGGCTGCGG 1272
QY 1578 CGATACGCTGGAATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1637
Db |||||
QY 1273 CGAGGCGCTGAGTGGAGATCCCTTCCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1332
QY 1638 CAG 1641
Db 1333 CGAG 1336

RESULT 2

US-09-252-991A-676/c

: Sequence 676, Application US/09252991A

: Patent No. 6551795

GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIORITY FILING DATE: 1999-02-18

; PRIORITY FILING DATE: 1999-02-18

; PRIORITY FILING DATE: 1998-02-18

; PRIORITY FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 676

; LENGTH: 1587

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

; US-09-252-991A-676

Query Match 30.1%; Score 503.6; DB 4; Length 1587;

Best Local Similarity 63.5%; Pred. No. 9.2e-133;

Matches 841; Conservative 0; Mismatches 464; Indels 19; Gaps 4;

QY 319 TTCGTGCTTTTGTAACTATCTGATGCGCTGCAAAATCGCGCTCGGATATGG-CCTT 377
Db |||||
QY 1420 TTCGTGCGCTTGGGCAACTGGATGATCCGCTGATGATCGCGCGCGGACATGCGCCCT 1361
QY 378 CCGCGTATGAACAACCTGCTGCTTCTGCTGCTTCAATGCGGTTACCGGATGCGCGCTGGC 437
Db |||||
QY 1360 CCGCGGATGAACAATTCAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1301
QY 438 TTCGCTGCTGCGACCGGCGGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 497
Db |||||
QY 1300 CACCTGCTTATGCTGCG 1253
QY 498 CCGCGCGCTGCGACCGCGGAGCTGCTTTCGATGCGACCTGCGGATTTTCGCGGTTCA 557
Db |||||
QY 1252 TGGCGGCTGCGACCGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1193
QY 558 CTTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 617
Db |||||
QY 1192 CTTGCGCGGATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1133
QY 618 CCG 677
Db |||||
QY 1132 CCG 1073
QY 678 GCGTGGCTGATCCTGCTGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCG 737
Db |||||
QY 1072 CCGCTTCTGCTGATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1013
QY 738 CGACCGTAACCTGCGGACGCTTCTTCAATCTGCTGCGCGCGGCTGACCCGATTTCTGTA 797
Db |||||
QY 1012 GGACATCCACTTCCG 953
QY 798 CCAACATCCTGCTGCTGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 857
Db |||||
QY 952 CCAGCAGTGTCTGCTGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 893
QY 858 TGGCATCATCAGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 917
Db |||||
QY 892 CGGTGCGCTGAGTGGCAATCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 833
QY 918 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 977
Db |||||
QY 832 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 773
QY 978 GTACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1037
Db |||||
QY 772 GTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 713
QY 1038 CCGGCTGCGGCGGCTTAAAGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1097
Db |||||
QY 712 CCGGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 653

Db 1126 ACCGAGTGTGTTCTCTGTTTCATGTCATCAAGTGCATCCGCGGCGCAAGCGGCGCCCT 1185
Qy 1555 CCGAACCCGTGGGCGGAATTCGCGGATACGCTGGAATGGAGCGTGCCTCATCCGCGCTCCG 1614
Db 1186 GCCAAGCCCTGGGAGC---GGCGGAGGGCTGGAGTGGAGCATCCCTCCGCGGCGCCC 1242
Qy 1615 GCCCAGACGTTGGAAGCGTGCCTCAAG 1641
Db 1243 TACCACACCTTCAGCACCCGCGCGAG 1269

RESULT 4
US-09-377-497-3
; Sequence 3, Application US/09377497
; Patent No. 6670119
; GENERAL INFORMATION:
; APPLICANT: YOSHIKAWA, YOSHIE
; APPLICANT: MUKAI, HIROYUKI
; APPLICANT: ASADA, KIYOZO
; APPLICANT: HINO, FUMISUGU
; APPLICANT: KATO, IKUNOSHIN
; TITLE OF INVENTION: CANCER-ASSOCIATED GENES
; FILE REFERENCE: 1422-389P
; CURRENT APPLICATION NUMBER: US/09/377,497
; CURRENT FILING DATE: 1999-08-20
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 3
; LENGTH: 1539
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: any n or xaa = unknown
US-09-377-497-3

Query Match 23.7%; Score 397.2; DB 4; Length 1539;
Best Local Similarity 58.1%; Pred. No. 1.3e-102;
Matches 743; Conservative 0; Mismatches 523; Indels 12; Gaps 2;

Qy 244 AACGGACACCTGGGAACGTCATGTTACCTACCATGGTATCTCATGATGTTCTTTGG 303
Db 148 AAGCACACATCACAACGTTATCGTCAGGCCCATGCAATTTGTAATATCTTCTCAT 207
Qy 304 GGTATCCCGCAATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 363
Db 208 GTAATACCATCATATAATCGGAGGCTTTGGCACTGACTAGTTCCCTTAATATCGGTGCC 267
Qy 364 CCGGATATGGCTTCGCGGTATGAACACCTGTGTTCTGGCTGTTCATTCGCGGTACC 423
Db 268 CCGGATATGGCTTCGCGGTATGAACACCTGTGTTCTGGCTGTTCATTCGCGGTACC 327
Qy 424 GCGATGGGCTGCTTGGCTGTGTCGACCGGCGGTGAGGTCAGTGGTTCGGGCGTT 483
Db 328 CTACTCTGCTGCACTCTGCTATAGTAGAGCGCGGAGCAGGAGTTGAACAGTC--- 384
Qy 484 GGTGGGTTCTGTACCCGCGGTGTGACCCCGGAGTGGCTATTCGATGAGACCTCGCG 543
Db 385 -----TACCCTCCCTTAGCAGGAACTACTCCACCTCGGAGCCTCGGTAGACCTAAC 438
Qy 544 ATTTTCGGGTTACATGTCGGTGCCTCTGATCATGGGCGGATCAACATGATCAG 603
Db 439 ATCTCTCTTACCTACCTAGAGTGTCTCTCTATCTTAGGGGCACTCAATTTCAATCA 498
Qy 604 ACTTCTTGAACATCGCGCCCGCGCATGAGCTGCACAAAGTCCCTTGTCTCGTGG 663
Db 499 ACAATATCAATATAAAACCCCTGCCATAACCCCAATACCAACGCGCTCTTCTGTGA 558
Qy 664 TCGATCTTATCAGCGTGGTGTGCTCTGCTGGGCTGCGGCTGCTGGTGTGCTGATC 723
Db 559 TCGGCTTAATCAGACAGTCTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 618
Qy 724 ACCATGCTGCTGACCGGACCGTAACTTCGCGCAGACCTTCTCAATCTCTGCTGGGCGGT 783

Db 619 ACTATACTACTAAACAGACCGCAACCTCAACACCACCTTCTTCGACCCCGCGGAGGGA 678
Qy 784 GACCCGATTCGTACCAACACATCTCTGGTGTCTTTGGGCAACCGGAAGTGTACATCATC 843
Db 679 GACCCGATTCGTACCAACACATCTCTGGTGTCTTTGGGCAACCGGAAGTGTATATCTT 738
Qy 844 ATTCGCGCGGCTTTGGCATCATCAGCCATCTGCTGTGACCTTCTC--GAAGGAGCG 900
Db 739 ATCTTACAGGCTTCGGAATATCTCCCAATATTAATTAATTAATTAATTAATTAATTA 798
Qy 901 GTCTTCGTTTACTTGGCGATGGTCTATGCAATGTGGCAATCGGTGTCTGGGCTTTGTC 960
Db 799 CCAITTTGATATAGGTATGGTCTGAGCTATGATATCAATTTGGCTTCTCTAGGTTTATC 858
Qy 961 GTCTGGGCGGACCATGATACACGTTGGTATGCTGACCCAGCAATCTCTGATCCCATG 1020
Db 859 GTGTGAGCACACCATATATTTACAGTAGGAATAGAGTAGACACAGAGCATATTTTACC 918
Qy 1021 CTGGCCACATGTTGATGGGTCGCGGTCGACCGGCAATTAAGATCTTCTGATGATCCACG 1080
Db 919 TCCGCTACCATATAATCATCGCTATATCCCAACCGGCTCAAGTATTTAGCTGACTCGCA 978
Qy 1081 ATGTGGGCGGCTCGGTTGAGTTCAAAATCGCGGATGCTCTGGGCTTTGGCTTTATGTC 1140
Db 979 CTCCAGGAGCAATATGAATGATCTGTGCACTGCTCTGAGCGCTTAGGATTCATCTTT 1038
Qy 1141 CTGTTTCCGTTGGTGGTGTGACCGGTATCGTGTGCGCAAGCGGTCTGGACCGTGCA 1200
Db 1039 CTTTTCACCGTAGGTGGCTGACTGCGATTTGATTAGCAAACTCATCACTAGACATCGTA 1098
Qy 1201 TATCAGACACCTTATAGTGTGGGCGCATCTTCCATTTATGTGATGCTGGTGGGATC 1260
Db 1099 CTACACGACACGTACTAGTGTAGCTCACTTCCATCTATCTCTCAATAGAGGTGTA 1158
Qy 1261 TTTCCGATCTTCCGCGTATCTACTTTTACATCGCAAGTCTTCGGGCGCGCTTTCCCG 1320
Db 1159 TTTGCCATCATAGAGGCTTCACTGATTTCCCTATTTCTCAGGCTACACCTAGAC 1218
Qy 1321 GAATGGCTGCAAGTGTCACTTCTGGAACCTTCTCATCGTTCGGAACGTCACGTTCTTC 1380
Db 1219 CAAACCTACGCCAAATCCATTTCCGTATCATATTCATCGCGCTAAATCTAACTTCTTC 1278
Qy 1381 CCGCAGCAGCTTCTTGGGAGCTGAGGTATCGCGCGGTACATCGACTATCCGAGCC 1440
Db 1279 CCACAAACCTTCTCGGCTTATCGGAATGCCCGAGCTTCTCGGACTACCCCGATGA 1338
Qy 1441 TTGCGCTGTGGAACAAAGTCTCTCTATGCTGGTTCCTGCGCTTCTGCGCTTCTGCTG 1500
Db 1339 TACACCATATAATATCCCTATCATCTGAGGCTCATTTCTTCTTAACAGCAGTAATA 1398
Qy 1501 TTCTTCATCTGATCTTT 1518
Db 1399 TTAATAATTTTCAGATT 1416

RESULT 5

US-08-219-842-1
; Sequence 1, Application US/08219842
; Patent No. 5563323
; GENERAL INFORMATION:
; APPLICANT: Parker, W. D.
; APPLICANT: Herinstdat, Corinna
; TITLE OF INVENTION: Diagnostic and Therapeutic Compositions
; TITLE OF INVENTION: for Alzheimer's Disease
; NUMBER OF SEQUENCES: 95
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; City: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:

NAME: Campbell, Cathryn A.
 REGISTRATION NUMBER: 31,815
 REFERENCE/DOCKET NUMBER: P-AG 9504
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 535-9001
 TELEFAX: (619) 535-8949
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1613 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-08-451-096-1

Query Match 23.7%; Score 397.2; DB 1; Length 1613;
 Best Local Similarity 58.1%; Pred. No. 1.4e-102;
 Matches 743; Conservative 0; Mismatches 523; Indels 12; Gaps 2;

244 AACGACACCTGTGGAAACGTCATGTTACCTTACCATGTTATCTTGATGATGTTCTTTG 303
 219 AACGACACATCTACAAGTATGTCACAGCCATGCAATTTGTAATATCTTTCTCA 278
 304 GGTATCCCGCAATTTGCGTGGTTTGGTAATCTGATGCGGCTGCAAAATCGGCGCT 363
 279 GTAATACCCATCAATATCGAGGCTTTGGCAACTGACTAGTTCCTCTAATATCGGTGC 338
 364 CCGGATATGGCTTCCCGCGTATGAACAACTGTGCTGCTGCTGCTGCTGCTGCTGCT 423
 339 CCGGATATGGCTTCCCGCGTATGAACAACTGACTAGTTCCTCTAATATCGGTGC 398
 424 CCGATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 483
 399 CTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 455
 484 GGTTGGGTTCTGTACCCGCGCTGTGACCCGCGGAGCTGCTGCTGCTGCTGCTGCTGCT 543
 456 -----TACCTCTCTAGCAGGAACTACTCCACCCCTGGAGCTCCGTAGACCTAAC 509
 544 ATTTTCGCGTTTCACTTGTGCGGTGCTCTCTGATCATGGGCGGATCAACATGATCAG 603
 510 ATCTCTCTTACACCTAGCAGTGTCTCTATCTTAGGGGCACTAATTTCAATCA 569
 604 ACCTCTTGAACATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 663
 570 ACAATATCAATAAAGCCCGTGCATACCAATACCAACGCGCTCTTCTGCTGA 629
 664 TCGATCTTTATACGCGTTGGCTGATCTCTGCGGCTGCGGCTGCTGCTGCTGCTGCTG 723
 630 TCGTCTTAATCACAGCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 689
 724 ACCATGCTGTGACGACGCTTACCTTGGCAGCACTCTTCAATCTCTGCGGCGGCT 783
 690 ACTATACTACTACAGACGCGCACTCTCAACACCACTCTCTGACCCCGCGGAGGA 749
 784 GACCCCAATCTGTACCAACACATCTGTGCTCTTTGGCACCCTGGAAGTGTACATATC 843
 750 GACCCCAATCTGTACCAACACATCTGTGCTCTTTGGCACCCTGGAAGTGTATATCT 809
 844 ATCTGCGCGCTTTGGCAATCATAGCCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 810 ATCTTACCAAGGCTCGGAATAATCTCCCATATTTGTAATCTACTCTCGGAAAGAA 869
 901 GCTCTGCTTACCTGCGGATGCTATGCAATGCTGCAATGCTGCTGCTGCTGCTGCTGCT 960
 870 CCAATTTGATATAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 929
 961 GTCTGGGCGCACCATGTATACCGCTTGGTATGTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 930 GTGTGAGCACACCATATATTTACATAGGAATAGACGTAGACACAGCAGATATTTTACC 989
 1021 CTGGCCACCATGCTGATGCGGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1080
 990 TCCGCTACCAATATCATGCTATCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1049

QY 1081 ATGTGGGGCGGCTCGGTGAGTTCAAAATCGCGGATCTCTGGGCGCTTGGCTTATGTT 1140
 DB 1050 CTCACCGGAAGCAATATGAATGATCTGCTGAGTGTCTGAGCCCTAGGATTCATCT 1109
 QY 1141 CTGTTACCGCTGGGTGTGACCGGTATCGTCTGGCCCAAGCGGTCTGGACCGTGA 1200
 DB 1110 CTTTTCACCGTAGTGGCTGACTGGCAATTTGATTAGCAAACTCATCACTAGACAT 1169
 QY 1201 TATCAGCACCTATTAGTGGTGGGCGACATTCATTTATGATGCTGCTGGTGGATC 1260
 DB 1170 CTACAGCACGCTACTAGCTTTAGCCCACTTCCACTATGCTCTATCAATAGGAGCT 1229
 QY 1261 TTTGCGATCTTCGCGGTATCTTACTTTTACATGCGCAAGTTCTCGGGCGGCTTTCCG 1320
 DB 1230 TTTGCCATCATAGGAGGCTTCAATTCACATGATTTCCCTATTTCTCAGGCTACCC 1289
 QY 1321 GAATGGGCTGCAAGCTGACATCTTGGACCTTTTCATCGTGGGAAGCTCACGTTTTC 1380
 DB 1290 CAACCTACGCGCAAAATCCATTTCACTATCATATTTCACTCGGCGTAAATCTAAC 1349
 QY 1381 CCGCAGCACTTCTCGGAGCGTCAAGGTATGCGGCGGTATACATCGACTATCCCGAG 1440
 DB 1350 CCACAGCACTTCTCGGCTATCGGATGCGGCGGATGCGGCGGCTACTCGGACTCC 1409
 QY 1441 TCGCGCTGTGGAAACAAAGTCTCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 1500
 DB 1410 TACACACATGAACATCTCTATCTGAGGCTCAITTCATTTCTTAACAGCAGTAATA 1469
 QY 1501 TTTCTCATCGTATCTTT 1518
 DB 1470 TTAATAATTTCAATGATT 1487

RESULT 7

US-08-810-599-1
 ; Sequence 1, Application US/08810599
 ; Patent No. 5976798
 ; GENERAL INFORMATION:
 ; APPLICANT: PARKER, W. Davis
 ; APPLICANT: HERENSTADT, Corinna
 ; APPLICANT: GHOSH, Soumitra S.
 ; APPLICANT: FAHY, Eoin
 ; TITLE OF INVENTION: Methods for Detecting Mitochondrial Mutations
 ; TITLE OF INVENTION: Diagnostic for Alzheimer's Disease and Methods for Determining
 ; TITLE OF INVENTION: of Mitochondrial Nucleic Acid
 ; NUMBER OF SEQUENCES: 82
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Kenyon & Kenyon
 ; STREET: 1025 Connecticut Avenue, N.W., Suite 600
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: US
 ; ZIP: 20036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.25" Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Wordperfect 6.1 for Windows
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/810,599
 ; FILING DATE: Concurrent Herewith
 ; CLASSIFICATION: 436
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/757,438
 ; FILING DATE: 27 No. 5976798 1996
 ; APPLICATION NUMBER: US 08/614,072
 ; FILING DATE: 12 Mar 1996
 ; APPLICATION NUMBER: US 08/536,036
 ; FILING DATE: 29 Sep 1995
 ; APPLICATION NUMBER: US 08/414,969
 ; FILING DATE: 31 Mar 1995
 ; APPLICATION NUMBER: US 08/413,740

FILING DATE: 30 Mar 1995
APPLICATION NUMBER: US 08/410,658
FILING DATE: 24 MARCH 1995
APPLICATION NUMBER: US 08/397,808
FILING DATE: 3 Mar 1995
APPLICATION NUMBER: US 08/219,842
FILING DATE: 30 MARCH 1994
ATTORNEY/AGENT INFORMATION:
NAME: Toffenetti, Judith L.
REGISTRATION NUMBER: 39,048
REFERENCE/DOCKET NUMBER: 2105/17
TELEPHONE: 202-429-1776
TELEFAX: 202-429-0796
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1613 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-810-599-1

Query Match 23.7%; Score 397.2; DB 2; Length 1613;
Best Local Similarity 59.1%; Pred. No. 1.4e-102;
Matches 743; Conservative 0; Mismatches 523; Indels 12; Gaps 2;

Qy 244 AAGGACACCTGTGGACGTCATGTTTACCTACCATGGTATCTGATGATGTTCTTTG 303
Db |||||
Qy 219 AAGACACATCTACACGTTATCGTCACAGCCATGCAATTTGTAATAATCTTTCATA 278
Db |||||
Qy 304 GGTATCCCGCATGTTTCGGTGGTTTGGTAACTATCTGATGCGCGTGCMAATCGGCGCT 363
Db |||||
Qy 279 GTAATACCATATAATCGGAGCTTTGGCACTGACTAGTTCCCTAATAATCGTGGC 338
Db |||||
Qy 364 CCGATATGCGCTTCCGGGTATGACACCTGCTCTTCTGGCTGTTCAATCGCGTACC 423
Db |||||
Qy 339 CCGATATGCGCTTCCCGCATAAACACATAAGCTTCTGACTCTTACCTCCCTCTCTC 398
Db |||||
Qy 424 GCGATGGCGGTGCTTCCGCTGTTTCGCACCGGGGGTGCAGGTGAGTTCGGCGGT 483
Db |||||
Qy 399 CTACTCTCTCGCATCTGCTATGTTGGAGCGCGGAGCAGGATGACATC--- 455
Db |||||
Qy 484 GGTGGGTTCTGTACCGCGCTGTGACCCGGAAGCTGGCTATTCGATGACCTCGCG 543
Db |||||
Qy 456 -----TACCCTCCCTTAGCAGGAACTACTCCACCCCTGGAGCCTCCGTAGACCTAAC 509
Db |||||
Qy 544 ATTTTCGCGTTCACTTGTGGGTGCTCTCTGATCATGGCGGATCAACATGATCAG 603
Db |||||
Qy 510 ATCTTCTCTTACACCTAGCAGGTGCTCTCTATCTTAGGGCCATCAATTTCAATCACA 569
Db |||||
Qy 604 ACCTTCTTGAACATGCGCGCCCGGCGATGACGCTGCACAAAGTGCGGTGTTCTCGTGG 663
Db |||||
Qy 570 ACAATTATCAATATAAACCCCTGCGCATAAACCAATACCAACGCGCTCTTCGCTGA 629
Db |||||
Qy 664 TCGATCTTATCAGGCTTGGCTGTGCTCTGCGCGCTCGGCTTCTGGCTGGTGAATC 723
Db |||||
Qy 630 TCCGTCCTATCAGCAGTCTTACTTCTCTATCTCTCCAGTCTAGCTCTGCTGCTC 689
Db |||||
Qy 724 ACATGCTGTGACCGACCTACTTTCGGCAAGCTCTTCAATCTCTGCTGGCGGGT 783
Db |||||
Qy 690 ACTATACTACTACAGACCGCAACCTCAACACCACTCTTCGACCCCGCGGAGGGA 749
Db |||||
Qy 784 GACCGATCTGTACCAACACATCTCTGTTGTTTGGGACCGGAAGTGTACATCATC 843
Db |||||
Qy 750 GACCCCATCTATACCAACACCTTCTGATTTTTCGTCACCTGAACTTATTTCT 809
Db |||||
Qy 844 ATTCTGCGCGCTTGGCATCATCAGCATGTCGTTGCGACCTCTC---GAAAGCGG 900
Db |||||
Qy 810 ATCTACAGGCTTCGGAATAATCTCCCATATTTGTAATCTACTCTCCGGAAGAAAGAA 869
Db |||||

Qy 901 GTCTCGGTTACCTGCGGATGGTCTATGCAATGTTGGCAATCGTGTCTTGGGCTTCTC 960
Db |||||
Qy 870 CCATTTGGATACATAGGTATGCTCTGAGCTATGATCAATTTGGATTTCTAGGTTTATC 929
Db |||||
Qy 961 GTCTGGCGGCACACATGTACACCGTTGCTATGCTGACCCACCAATCTCTACTTCTATG 1020
Db |||||
Qy 930 GTGTGAGCACACCATATATTTACAGTAGGAATAGACGTAGACACAGGACATATTTTACC 989
Db |||||
Qy 1021 CTGCGCACCATGATGTCGCGGTGCGGACCGGCAATTAAGATCTTCTGTTGATCGCCACG 1080
Db |||||
Qy 990 TCGGCTACCATATATCATGCTATCCCAACCGGCTCAAAGTATTTAGCTGACTCGCCACA 1049
Db |||||
Qy 1081 ATGTGGGGCGGCTGCTGAGTTCAAATCGCGATGCTCTGGGCTTTGGCTTTATGTTTC 1140
Db |||||
Qy 1050 CTCACGGAAGCAATATGAATGATCTGCTGAGTGTCTGAGCCCTAGGATTTCTATCTTT 1109
Db |||||
Qy 1141 CTGTTCACTGCTGGTGTGACCGGTATCGTGTGGCCCAAGCGGCTCTGGACCGTGA 1200
Db |||||
Qy 1110 CTTTTCACGCTAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1169
Db |||||
Qy 1201 TATCAGCACCTATTACGTTGGTGGCGGCTTCCATTTATGATGCTGCTGGTGGGATC 1260
Db |||||
Qy 1170 CTACAGCACGCTACTGCTGTTAGCCCACTTCCACTATGCTCTATCAATAGGAGCTGTA 1229
Db |||||
Qy 1261 TTTGCGATCTTCCCGGTATCTATTTTACATGCCGAAGTCTTCCGGCGGCGCTTTCCCG 1320
Db |||||
Qy 1230 TTTGCGCATCATAGGCTTCTTCACTGATTTCCCTATTTCTCAGGCTACACCCCTAGAC 1289
Db |||||
Qy 1321 GAATGGCTGCAAGCTGCTGACCTTCTGACCTTCTTCACTGGTGGCGGCTGACGCTTCTTC 1380
Db |||||
Qy 1290 CAACTACGCCAAATCCATTTCACTATCATATTCGCGGTAAATCTAACTTTCTTCTC 1349
Db |||||
Qy 1381 CCGCAGCATCTTCTGGAGCTCAGGTATGCGCGCGGCTTACATCGACTATCCGGAAGC 1440
Db |||||
Qy 1350 CCAACACATTTCTGGCTATCCGGAATGCCCGAGTTACTCGGACTACCCCGATGCA 1409
Db |||||
Qy 1441 TTTGCGCTGGAACAACTCTGCTCTATGTTGGTGGTCTGCGCTTCCGCTGCTGCTGCTG 1500
Db |||||
Qy 1410 TACACACATGAACATCTCTATCATCTGTAGGCTCAATCTTCTTAACAGCAGTAATA 1469
Db |||||
Qy 1501 TTCTTCATCTGATCTTT 1518
Db |||||
Qy 1470 TTAATAATTTCTGATT 1487
Db |||||

RESULT 8

US-09-525-906-1
; Sequence 1, Application US/09525906
; Patent No. 5605433
; GENERAL INFORMATION:
; APPLICANT: Jen, Jen
; APPLICANT: Sidransky, David
; APPLICANT: Kinzler, Kenneth
; APPLICANT: Vogelstein, Bert
; APPLICANT: Fliss, Makiko
; APPLICANT: Polyak, Kornelia
; TITLE OF INVENTION: Mitochondrial Dosimeter
; FILE REFERENCE: 1107.85815
; CURRENT APPLICATION NUMBER: US/09/525,906
; CURRENT FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 09/377,856
; PRIOR FILING DATE: 1999-08-20
; PRIOR APPLICATION NUMBER: 60/097,307
; PRIOR FILING DATE: 1998-08-20
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 16568
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-525-906-1

Query Match

23.7%; Score 397.2; DB 4; Length 16568;

QY 364 CCGGATATGGCCTTCCCGGATGAAACAAACCTGTGCTTCTGGCTGTTCAITTCGCGGTACC 423
Db 6171 CCCGATATGGCGTTTCCCGGCAATAAACACATAAAGCTTCTGACTCTTACCTCCCTCTCTC 6230
QY 424 GCGATGGGCGTGGCTTGGCTGTTCGACCGGCGGTGACGGTCAGCTGGGTTCGGGCGTT 483
Db 6231 CTACTCTCTGCTGCACTCTGCTATAGTGGAGCGGAGGACAGAGTGTGAACAGTC--- 6287
QY 484 GGTGGGTTCGTATACCGGCGGTGTGACCGCGGAGCTGGCTATTCGATGGACCTCGCG 543
Db 6288 -----TACCTCTCTTAGCAGGAAGTACTCCACCTCGGAGCGCTCGGTAGACCTAAC 6341
QY 544 ATTTTCGGGTTCACTTGCCTGGGTGCTCTCGATCATGGGCGGATCAACATGATACAG 603
Db 6342 ATCTTCTCTTACACTAGAGGTGTCTCTCTATCTTAGGGCCATCAATTTCAACACA 6401
QY 604 ACCTTCTTGAACATGCGCGCCCGGCATGACGTGCAAAAGTCCGCTGTCTTCGTGG 663
Db 6402 ACAATTTATCAATAAAACCCCTGCCATAACCCCAATACCAACGCCCTCTTCTGTGA 6461
QY 664 TCGATCTTTATCAGCGCTTGGCTGATCTCTGCTGGCGCTGCCGTTCTGGCTGGTCAATC 723
Db 6462 TCCGTCCTAATCACAGCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 6521
QY 724 ACCATGCTGCTGACCGACCGGTAACTTTCGGCACGACCTTCTTCAATCTCTGCTGGCGG 783
Db 6522 ACTATACTACTAACAGACCGCAACCTCAACACCACTTCTTCGACCCCGCGGAGGAGGA 6581
QY 784 GACCGGATCTGTACCAACACATCTCTGTGGTCTTTTGGGACCGGAGTGTATCATCATC 843
Db 6582 GACCCCACTTATACCAACACCTATCTCTGATTTTGGGTACCCCTGAAGTTTATATTT 6641
QY 844 ATTCTCCCGGCTTGGCATATCAGCATGTGCTGCTGACCTTCTC---GAAAGACCG 900
Db 6642 ATCTTACGAGGCTTCGGAATATCTCCCATATTTGAATCTTACTCTCCGGAATAAGAA 6701
QY 901 GTCTTCGGTTACCTGCGGATGCTTANGCAATGGTGGCAATCGGTTCCTGGCTTTGTTC 960
Db 6702 CCATTTGGATACATAGTATGCTGAGCTATGATATCAATTTGGCTTCTCTAGGTTTATC 6761
QY 961 GTCTGGGCGCACCATGTATACCGTTGATGCTGCTGACCGACCAATCTTACTTCATG 1020
Db 6762 GTGTGAGCACACCATATATTTACGTAGGAATAGAGTATAGACATATTTTACCC 6821
QY 1021 CTGGCCACCATGTGATCGCGGTGCGGACCGCATTAAGATCTTCTGTTGATCCGACG 1080
Db 6822 TCCGCTACCATATCATCGCTATCCCAACCGGCTCAAAAGTATTTAGCTGACTGCCACA 6881
QY 1081 ATGTGGGCGGCTCGGTGAGTTCAATCGCGATGCTCTGGGCTTTGGCTTTATGTC 1140
Db 6882 CTCACGGAAGCAATATGAATGATCTGTGCACTGCTCTGAGCCCTAGGATTCATCTTT 6941
QY 1141 CTGTTTACCGTGGGTGTGACCGGTATCTGCTGGCCCAAGCGGTCTGGACCGTGCA 1200
Db 6942 CTTTTCACCGTAGTGGCTGACTGCAATTTATAGCAAACTCATCATAGACATCGTA 7001
QY 1201 TATCAGACACCTATTACGTGGTGGGCACTTCCATATATGATGATGCTGGGTGGGATC 1260
Db 7002 CTACAGACACGTAAGTGTAGGCCACTTCCACTATGCTCTATCAATAGAGCTGTA 7061
QY 1261 TTTTGGATCTTCCCGGTATCTACTTTTACATGCCGAAGTCTTCCGGCGCGCTTTCCCG 1320
Db 7062 TTTTGGCATATAGAGGCTTCACTGATTTTCCCTATTTCTAGGCTTACACCTAGAC 7121
QY 1321 GAATGGGCTGCAAAAGTGTGCTTCTGGAATCTTCTCATCGGTGCGAACGTCAAGTTCTTC 1380
Db 7122 CAAACCTTACGCCAAATCCATTTTCACTATCATATTTTATCGGCGTAAATCTAACTTTCTTC 7181
QY 1381 CCGAGACACTTCTGGGACGTGAGGTATGCGCGCGGTGTACATCACTATCCGAGCC 1440
Db 7182 CCACAAACACTTCTCGGCTTATCCGGAATGCGCGGAGCTTACTCGGACTACCCCGATGCA 7241

QY 1441 TTCGGCTGTGGAAAGTCTCGTCTATGTTGGTCTTCTCGCTCGCTCGCTCGTCTG 1500
Db 7242 TACACACATGAACATCTCTATCATCTGAGGTCTCATTTCTCTTAACAGCAGTAATA 7301
QY 1501 TTTCTCATCTGATCTTT 1518
Db 7302 TTAATATTTTCATGATT 7319

RESULT 10
US-09-377-856-1
; Sequence 1, Application US/09377856
; Patent No. 6344322
; GENERAL INFORMATION:
; APPLICANT: Polyak, Kornelia
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Subtle Mitochondrial Mutations as Tumor
; FILE REFERENCE: 1107.82346
; CURRENT APPLICATION NUMBER: US/09/377,856
; CURRENT FILING DATE: 1999-08-20
; PRIOR APPLICATION NUMBER: 60/097,307
; PRIOR FILING DATE: 1998-08-20
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 16569
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-377-856-1

Query Match 23.7%; Score 397.2; DB 3; Length 16569;
Best Local Similarity 58.1%; Pred. No. 4.3e-102;
Matches 743; Conservative 0; Mismatches 523; Indels 12; Gaps 2;

QY 244 AACGGACACCTGTGGAACGTCACTTACCTACCATGTTCTGATGATGTTCTTTGTG 303
Db 6051 AACGACACATCTCAACAGTTATCTGACAGCCCATGATTTGTAAATCTTTCTATA 6110
QY 304 GGTATCCCCCGATTGTTGGTGGTTTTGGTAACCTATCTGATGCGCTGCAAAATCGGCGT 363
Db 6111 GTATATCCCATCAATATCGGAGGCTTTGGCAACTGACTAGTTCCCTTAATAATCGGTGCC 6170
QY 364 CCGATATGGCTTCCCGGTATGAACAACTGTCTGTTCTGGCTGTTTCATGCGGTACC 423
Db 6171 CCCGATATGGGTTTCCCGCATAAACATAGCTTCTGACTCTTACCTCCCTCTCTC 6230
QY 424 GCGATGGCGGTGCTGCTGTTCCGACCGGCGGTGACGTCAGCTGGGTTCCGGGCTT 483
Db 6231 CTACTCTGCTCGCATCTGCTATAGTGGAGCGCGAGACAGAGTGTGAACAGTC--- 6287
QY 484 GGTGGGTTCGTATACCGCGCGCTGTGACCCGCGAGCTGGCTATTCGATGGACCTCGCG 543
Db 6288 -----TACCTCTCTTAGCAGGAAGTACTCTCCACCGCTCGAGCTCCGTAGACCTAAC 6341
QY 544 ATTTTCGGGTTCACTTGTGGGTGCTCTCTGATCATGGGCGGATCAACATGATACAG 603
Db 6342 ATCTTCTCTTACACTAGAGGTGTCTCTCTATCTTAGGGCCATCAATTTCAACACA 6401
QY 604 ACCTTCTTGAACATGCGCGCCCGGCATGACGTGCAAAAGTGGCTTCTTCTGCTGG 663
Db 6402 ACAATTTATCAATAAAACCCCTGCCATAACCCCAATACCAACGCCCTCTTCTGTGA 6461
QY 664 TCGATCTTTATCAGCGGTGCTGATCTCTGCTGGCGCTGCCGTTCTGGCTGGTCAATC 723
Db 6462 TCCGTCCTAATCACAGCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 6521
QY 724 ACCATGCTGCTGACCGACCGGTAACTTTCGGCACGACCTTCTTCAATCTCTGCTGGCGG 783
Db 6522 ACTATACTACTAACAGACCGCAACCTCAACACCACTTCTTCGACCCCGCGGAGGAGGA 6581
QY 784 GACCGGATCTGTACCAACACATCTCTGTGGTCTTTTGGGACCGGAGTGTATCATCATC 843

[illegible]

RESULT 11

```

US-09-302-681-2
; Sequence 2, Application US/09302681
; Patent No. 6441149
; GENERAL INFORMATION:
; APPLICANT: Herrstadt, Corrina
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Clevenger, William
; APPLICANT: Fahy, Eoin F.
; APPLICANT: Davis, Robert E.
; TITLE OF INVENTION: DIAGNOSTIC METHOD BASED ON
; TITLE OF INVENTION: QUANTIFICATION OF EXTRAMITOCHONDRIAL DNA
; FILE REFERENCE: 660088.416C1
; CURRENT APPLICATION NUMBER: US/09/302,681
; CURRENT FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 16569
; TYPE: DNA
; ORGANISM: Homo sapien

```


QY 1381 CCGCAGACATCTCTGGGACGCTCAGGATATGCGCGCGGTTTACATCGACTATCCGAGCC 1440
Db 7182 CCACACATCTTCTGGGCTATCCGGAATGCCGAGCTTACTCGGACTACCCCGGATGA 7241
QY 1441 TTCCGCGTGTGGAAACAAAGTCTCGTCTATGGTGGCTTCTGGCCCTTCGCTCGTTCCTG 1500
Db 7242 TACACACATGAACATCCATCATCTGTAGGCTCATTCATTTCTCTACAGCAGTAATA 7301
QY 1501 TTCTTCATCGTATCTTT 1518
Db 7302 TTAATAATTTTCATGATT 7319

RESULT 13
US-10-053-611-1
; Sequence 1, Application US/10053611
; Patent No. 6750021
; GENERAL INFORMATION:
; APPLICANT: Polyak, Kornelia
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Subtle Mitochondrial Mutations as Tumor
; TITLE OF INVENTION: Markers
; FILE REFERENCE: 1107.82346
; CURRENT APPLICATION NUMBER: US/10/053,611
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: US/09/377,856
; PRIOR FILING DATE: 1999-08-20
; PRIOR APPLICATION NUMBER: 60/097,307
; PRIOR FILING DATE: 1998-08-20
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 16569
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-053-611-1

Query Match 23.7%; Score 397.2; DB 4; Length 16569;
Best Local Similarity 58.1%; Pred. No. 4.3e-102;
Matches 743; Conservative 0; Mismatches 523; Indels 12; Gaps 2;

QY 244 AAGGACACCTGTGGACGTCATGTTTACCTACCGATTTCTGATGATGTTCTTTG 303
Db 6051 AAGACACATCTACACGTTTTCGTACAGCCATGCTTTGTAAATCTTCTTCA 6110
QY 304 GGATATCCCGCATGTTTGGTGGTTTGGTAACTATCTGATGCGGTGCAAAATCGCGCT 363
Db 6111 GTAATACCATCAATATCGGAGGCTTGGCACTGACTAGTTCCCTTAATAATCGGTCC 6170
QY 364 CCGGATATGGCTTCCCGGTATGAACAACTGCTCTTGGTGTTCATTGCGGTACC 423
Db 6171 CCGGATATGGCTTCCCGGTATGAACAACTGCTCTTGGTGTTCATTGCGGTACC 6230
QY 424 GCGATGGCGTGGCTTGGCTGTTGCGACCGCGGTGACGCTCAGCTGGGTTCCGGCGTT 483
Db 6231 CTACTCTGCTCGCATCTGCTATAGTGGAGCGGAGCAGGATGAACAGTC--- 6287
QY 484 GTTTGGGTTCTGATCCCGCGCTGTGACCCGGAAGCTGGCTATTCGATGGACCTCGCG 543
Db 6288 -----TACCCTCCCTTAGCAGGAACTACTCCACCCCTGGAGCCCTCGTAGACCTAAC 6341
QY 544 ATTTTCGCGTTTCACTTGTGGGTGCTCTCTCGATCATGGGCGCATCAACATGATCA 603
Db 6342 ATCTTCTCTTACCTAGCAGTGTCTCTCTATCTTAGGGGCTCATATTTTATCA 6401
QY 604 ACTTCTTGAACATGGCGGCGCGGATGACCTGCAAAAGTGGCTGTGTTCTGGTG 663
Db 6402 ACAATTATCAATAAACCCTTGCATAAACCAATACCAACGCGCTCTTCTGCTCA 6461
QY 664 TCGATCTTTATCAGGCTTGGCTGATCCCTGCTGGCGCTCGGCTTCTGCTGCTGCTCAATC 723
Db 6462 TCGTCTCTATCAGCAGTCTTACTTCTCTCTATCTCTCCAGTCTCTAGCTCTGGCATC 6521

QY 724 ACCATGCTGCTGACCGACCGTAACCTTGGGACGACCTTCTTCAATCTCTGCTGGCGCGGT 783
Db 6522 ACTATACTACTAAACAGACCGCACTCAACACCACTTCTTCGACCCCGCGGAGGAGGA 6581
QY 784 GACCCGATTTCTGTACCAACACATCTGTTGGTTCCTTGGGACCGCGGAGTGTACATCATC 843
Db 6582 GACCCATTTCTATACCAACACCTATTCTGATTTTTCGTCACCTGAAGTTTATATCTT 6641
QY 844 ATTCTGCGCGGTTTGGCATCATCAGCAGCATCTGCTGCGACCTTCTC---GAAAAAGCCG 900
Db 6642 ATCTTACCGAGCTTCGGAATAATCTCCCATATTCTACTCTCGGAAAAAGAA 6701
QY 901 GTCTTCGGTTTACCTGCGATGCTCTATGCAATGTTGCAATCGGTGTTCTGGGCTTTGTC 960
Db 6702 CCAATTTGATACATAGGTATGCTCTGAGCTATGATATCAATTTGCTTCTAGGTTTATC 6761
QY 961 GTCTGGCGCACCATATGATACACCGTTGGTATGCTGCGTACCCAGCAATCTACTTCATG 1020
Db 6762 GTGTGAGCACACCATATATTTACAGTAGGATAGACGTAGACACACGAGCATATTTTACC 6821
QY 1021 CTGGCCACCATGTTGATCGCGTGGTGGCGGACCGCATTAAGATCTTCTCGTGGATCGCCACG 1080
Db 6822 TCCGCTACCATTAATCATCTGCTATCCACCGCGCTCAAAGTATTTAGCTGACTCGGCACA 6881
QY 1081 ATGTGGGCGGCTCGGTTGAGTTTCAAAATCGCCGATGCTCTGGGCTTTTGGCTTTATGTT 1140
Db 6882 CTCCACGAGCAATATGAATGATCTGCTGACGCTCTCTGAGCCCTTAGGATTCATCTTT 6941
QY 1141 CTGTTACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1200
Db 6942 CTTTTCACCGTAGTGGCTGACTGGCATTTGATTTAGCAAACTCATCACTAGACATCGTA 7001
QY 1201 TATCAGCACACTATTAGTGGTGGCGCACCTTCATTATGATGTCGCTGGGTGGCGATC 1260
Db 7002 CTACAGCACGTTACTAGTGGTGGCGCACCTTCATTATGATGTCGCTGGGTGGCGATC 7061
QY 1261 TTTGCGATTTTGGCGGTAICTATTTTATACGCGGAAGTTCTGGGCGCGGCTTTCCCG 1320
Db 7062 TTTGCGATCATAGGAGGCTTCAATTCATGATTTTCCCTATTTCTCAGGCTACACCTAGAC 7121
QY 1321 GAATGGGCTGCAAGCTGCACTTCTGGACCTTCTTCATCGTGGCGAACGTCACGTTCTTC 1380
Db 7122 CAACCTACGCGCAAAATCCATTTCACTATATATTCATCGCGGTAATCTAACTTTCTTC 7181
QY 1381 CCGCAGACATCTCTGGGACGCTAGGCTATGCGCGCGCTTACATCGACTATCCGGAAGCC 1440
Db 7182 CCACACACATTTCTCGGCTATCCGGAATGCCGAGTTTACTCGGACTACCCCGATGCA 7241
QY 1441 TTCCGCTGTGGACAAAGTCTGCTCTATGTTGGTTCCTGCGCTTCTGCGCTTCTGCTGCTG 1500
Db 7242 TACACCATGAAACATCTTATCATCTGTAGGCTCATTCATTTCTTCTAACAGCAGTAATA 7301
QY 1501 TTCTTCATCGTATCTTT 1518
Db 7302 TTAATAATTTTCATGATT 7319

RESULT 14
US-09-302-681-3
; Sequence 3, Application US/09302681
; Patent No. 6441149
; GENERAL INFORMATION:
; APPLICANT: Herrnstadt, Corrina
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Clevenger, William
; APPLICANT: Fahy, Boi F.
; APPLICANT: Davis, Robert E.
; TITLE OF INVENTION: DIAGNOSTIC METHOD BASED ON
; FILE REFERENCE: 650088.416C1
; CURRENT APPLICATION NUMBER: US/09/302,681
; CURRENT FILING DATE: 1999-04-30

```
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 6691
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(6691)
; OTHER INFORMATION: n = A,T,C or G
US-09-302-681-3

Query Match      23.6%; Score 395.2; DB 4; Length 6691;
Best Local Similarity 57.7%; Pred. No. 1e-101;
Matches 738; Conservative 3; Mismatches 525; Indels 12; Gaps 2;

QY 244 AACGACACCTGTGACGTCATGTTACCTACCATGGTATCTCATGATGTTCTTTTGG 303
DB 2582 AACGACACCATCTACACGTTATCTGTCACAGCCCATGATTTGTAATATCTTTCATA 2641
QY 304 GGTATCCCGCATTTGTCGGTGGTTTGGTAACTATCTGATCCGCTCAAAATCGCGCT 363
DB 2642 GTAATACCCATCAATAATCGAGGYTTTGGCAASTGACTAGTTCCCTAATAATCGGTGCC 2701
QY 364 CGGATATGGCTTCCCGGTATGAACAACCTGTGTTCTGCTGTTCTTCAATGCGGTACC 423
DB 2702 CCGGATATGGGTTTCCCGCGCATAAACAACAAGCTTNGACTCTTACCCGCCCTCTCTC 2761
QY 424 CGGATGGGCTGGCTTCGGCTGTTCGACCGCGGGTGACCGTCAGTCGGTTCGGGCGGT 483
DB 2762 NTACTCCTGTNGCATCTCTATATGAGGAGCGGGG-----CAGGACAGGTTGA 2812
QY 484 GGTGGGTTCTGTACCCCGCGGTGTGCGACCCCGGAAGCTGGCTATTCGATGACCTCGG 543
DB 2813 ACAGTNTACCCCTCTTGGCAGGAACTACTCCACCCTGGAGCCTCCGTAGACCTAAC 2872
QY 544 ATTTTCGGGTTACACTGTGCGGTGCTCTCGATCATCGGCGCATCAACATGATCAG 603
DB 2873 ATCTTCTCTTACACCTAGCAGGTATCTCTTAICTTAGGACCATCAATTTTCATCA 2932
QY 604 ACCTTCTGAACATGCGCGCCCGCCGATGACGCTGACAAAGTGGCGTTTTCGTGG 663
DB 2933 ACAATPATTAAATAAAACCCCTGCGCATAAACCAATACAAACGCCCTTTTCGTGTA 2992
QY 664 TCGATCTTTATCAGGCTTGGCTGATCCTGCTGGCGTCCGCTGCTGGCTGCTGCTCAATC 723
DB 2993 TCGTCTCTAATCAGCAGCTTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3052
QY 724 ACCATGCTGTCACCGACGCTAACTTCGGCAGCACCTTCTTCAATCCTGCTGCGCGGT 783
DB 3053 ACTATACTACTACAGACCGTAACCTCAACACCACTTCTTCGACCCAGCGAGGAGA 3112
QY 784 GACCGATTTCTPACCAACACATCTCTGTGGTTCTTTGGGACCCGGAAGTGTACATCATC 843
DB 3113 GACCCCATTTCTATACCAACACCTATTCTGATTTTTCGGTCACCCCTGAAGTTTATATCTC 3172
QY 844 ATTCTCGCGGCTTGGCATCATCAGCCATGCTGCTGACCTTCTC---GAAAGCGG 900
DB 3173 ATCTTACAGGCTTCGATTAATCTCCATATTGTAACCTACTCTCCGGAATAAGAA 3232
QY 901 GTCTTCGGTTTACCTGCGGATGGTCTATGCAATGGTGGCAATCGGTTTCTGGGCTTTGTC 960
DB 3233 CCATTTGGATACATAGTATGCTGAGCTATGATATCAATGGCTTCTAGGTTTATC 3292
QY 961 GTCTGGGCGACCATGTACACCGTGGTATGCTGCTGACCGACCAATCTCTACTCATG 1020
DB 3293 GTGTGAGCACCATATATTTCAGTAGGATAGACGTAGACACAGACGATATTTCACC 3352
QY 1021 CTGCGCACCATGTTGTCGGGTGCGGACCGGATTAAGATCTTCTCGTGATCGCCAG 1080
DB 3353 TCCGCTACCATATATCTGCTATCCCGACCGGCTCAAGATTTTAGCTGACTCGCCACA 3412
QY 1081 ATGTGGGCGGCTCGGTTGAGTTCAATCGCCGATGCTCTGGGCCCTTGGCTTTATGTTTC 1140
```

RESULT 15

```
US-09-097-889-1
; Sequence 1, Application US/09097889
; Patent No. 6218117
; GENERAL INFORMATION:
; APPLICANT: Herxstadt, Corrina
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Davis, Robert E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING
; TITLE OF INVENTION: AGENTS THAT QUANTITATIVELY ALTER DETECTABLE
; TITLE OF INVENTION: EXTRAMITOCHONDRIAL DNA: MITOCHONDRIAL DNA RATIOS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/097,889
; FILING DATE: 15-JUN-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rosenman P.D., Stephen J.
; REGISTRATION NUMBER: 43,058
; REFERENCE/DOCKET NUMBER: 660088.417
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6744 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

US-09-097-889-1

Query Match 23.2%; Score 388.2; DB 3; Length 6744;
Best Local Similarity 57.1%; Pred. No. 9.9e-100;
Matches 730; Conservative 19; Mismatches 517; Indels 13; Gaps 3;

```
QY 244 AACGGACACCTGTGGAACGTCATCGTTACCTACCATGTTATCTTCATGATGATGTTCTTTGIG 303
DB 2625 AACGACACACATCTACACAGTATCTGTCACAGCCATGCAATTTGTAATAATCTTTTCATA 2684
QY 304 GGTATCCCCCGAATGTTTCGGTGGTTTTGGTAACATCTGATGCGCGTGCAGAAATCGGCGT 363
DB 2685 GTAATACCCATCAATAATCGGAGGTTTGGCAACTGACTAGTTCCTCCCTAATAATCGGTGCC 2744
QY 364 CCGGATATGGCTTCCCGCGGTATGAACAACTGTCGTTCTGGGTGTTTCATTGCGCGTACG 423
DB 2745 CCGATATGGGTTTCCCGCATAAACAATAAGCTTTGACTTTTACCTCCCTCTCTC 2804
QY 424 GCGATGGCGGTGCTTGGCTTTGCGACCGGGGTGACGTCAGTGGGTTTCGGGCGTT 483
DB 2805 MTACTCTCTGTYGCATCTGCTATAGTGGAGCGCGMGACGAAACAGGTTGAACAGTM--- 2861
QY 484 GGTGGGTTCTGTACCGCGCGCTGTGACCGCGAGCTGGCTATTGCGAGCTCGG 543
DB 2862 -----TACCTCCCTTGGAGGAACTACTCCACCMGTGGAGCTCCGTTAGACSTAACC 2915
QY 544 ATTTTCGGGTTCACTTGTGCGGTGCTTCG-TGGATCATGGGCGGATCAACATGATCAC 602
DB 2916 ATCTTSCCTTACACYTAGCAGGTCTCTCTCTTAICTTAGGGGCCATCAATTTCAATCAC 2975
QY 603 GACCTTTTGAACATGGCGCGCCCGCGCATGACGTCGCACAAAGTGGGTTTCTCGTG 662
DB 2976 AACAAATATYAATAAATAAACCCTCGCATAAACCCCAATACCAAAACGCCCTTCTGCTG 3035
QY 663 GTCGATCTTTATACGGCTTGGCTGATCTGCTGGCGTCCCGGTTCTGGCTGGTGCAAT 722
DB 3036 ATCCGTCCTAATCACACAGTCVCTACTTCTCTCTCTCTCCAGTCTAGCYGCTGGCAT 3095
QY 723 CACATGCTGCTGACGACCGTAATCTGCGGACGACCTTCTCAATCTGCTGGCGGG 782
DB 3096 CACTATACTACTAACAGACCGYAMCYTCAACACCACTTCTTCTGACCCCTGAGGAGG 3155
QY 783 TGACCCGATCTGTACCAACACATCTGTGTTCTTTGGGACCCCGGAAAGTGTACATCAT 842
DB 3156 AGACCCCATTTCTATACCAACACCTATTTCTGATTTTGGTTCACCTGAGTTTATTTCT 3215
QY 843 CATCTCCCGGGTTTGGCATATGACCCATGTGCTGCGACCTTCTCG---AAAAAGCC 899
DB 3216 YATCCTACACAGGTTTCGGAATAATCTCCCATATTTGTAATCTACTCCGGMAAAAAAGA 3275
QY 900 GGCTTCGGTTACTCCGCTGCTATGCAATGTTGGGATCGGTTCTGGGCTTTGT 959
DB 3276 ACATTTGGATACATAGGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3335
QY 960 CGTCTGGGCGCACCATATGACACCTTGGTATGCTGCTGACCCAGCAATCTCTTCTCAT 1019
DB 3336 CGTGTGAGCACACCATATATTTACAGTAGGAATAGACGTAGACACAGGACATATTTTAC 3395
QY 1020 GCTGGCCACCATGTTGATCGGTTGGGACCGGATTAAGATCTTCTGTTGATCGCCAC 1079
DB 3396 CTCGGTTACCATATATCATGCTATCTCCACCGGCGTCAAAGTATTTAGCTGACTCGCCAC 3455
QY 1080 GATGTGGGGGGCTCGGTTGAGTTTCAATCGCCGATGCTCTGGGCGCTTTGGGCTTTATGTT 1139
DB 3456 ACTCCAGGAAGCAATATGAATGATCTGCTGAGTGTCTGTAGCCCTAGGATTTATTTT 3515
QY 1140 CTTGTTACCGTGGGTGTGACCGGTATCGGTGCTGGGCCAAGCGGGTCTGGAACGGTGC 1199
DB 3516 TCTTTTACCGTAGGTGGCTGACTGGCATTTGTTATGCAAACTCATCTACTAGACATCGT 3575
QY 1200 ATATCAGCACCATTTATGCTGTTGGGACCTTCCATTTATGATGCTGCTGGTGGCAT 1259
DB 3576 ACTACAGCACAGTACTACGTTGTAGCCCACTTCCACTATGCTCTATCAATAGGAGCTGT 3635
```

Search completed: October 22, 2004, 12:24:08
Job time : 181 secs

```
QY 1260 CTTTCGGATCTTCCCGGTATCTACTTTTATGTCGCGAAGTTCTTCGGGCCCGGCTTTCCC 1319
DB 3636 ATTTGCCATCATAGGAGGCTTTCATCTACTGATTTTCCCTATTTCTCAGGGTACACCCCTAGA 3695
QY 1320 GGAATGGGCTGCAAAAGCTGCACCTTCTGACCTTCTCATCGGTGCGAAGCTCAGTTCTT 1379
DB 3696 CCAAACTTACGCAAAATCGATTTTCGTATCATATTCATCGCGCTAAATCTAACTTTCTT 3755
QY 1380 CCGGACGACTTCTCTGGGACGTCAGGTATGCGCGCGGTATCATCTGACTATCCGGAAGC 1439
DB 3756 CCCACAAACACTTCTCTCGGCTATCCGGAATGCCCGGACGTTTACTCGGACTAYCCCGATGC 3815
QY 1440 CTTCCGCGCTGCGAACAAGTCTGCTCTATGTTGCGGTTCTGCGCTTCTGCGCTCGTTCT 1499
DB 3816 ATACACCATGAAYATCCTATCATCTGAGGTCTCATTTCTCTAACAGCAGTAAT 3875
QY 1500 GTTCTTCATCGTATCTTT 1518
DB 3876 ATTAATAATTTTCATAAT 3894
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 22, 2004, 11:44:00 ; Search time 856 Seconds
(without alignments)
10020.378 Million cell updates/sec

Title: US-09-712-768A-1

Perfect score: 1674

Sequence: 1 atggcagacgcgcattca.....acaagcatccctgcactaa 1674

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3407233 seqs, 2561960514 residues

Total number of hits satisfying chosen parameters: 6814466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

```

1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	515.6	30.8	1593	15	US-10-127-032-32
2	515.6	30.8	1593	16	Sequence 32, Appl
3	406.2	24.3	1629	16	Sequence 9, Appl
4	398.8	23.8	16559	16	Sequence 25546, A
5	398.8	23.8	16559	16	Sequence 150, App
6	398.8	23.8	16559	16	Sequence 145, App
7	398.8	23.8	16559	16	Sequence 144, App
8	398.8	23.8	16559	16	Sequence 181, App
9	398.8	23.8	16559	16	Sequence 440, App
10	398.8	23.8	16559	16	Sequence 469, App
11	398.8	23.8	16559	16	Sequence 474, App
12	398.8	23.8	16559	16	Sequence 476, App
					Sequence 501, App

Query Match 30.8%; Score 515.6; DB 15; Length 1593;

Best Local Similarity 63.0%; Pred. No. 7e-146;

Matches 854; Conservative 0; Mismatches 484; Indels 18; Gaps 3;

Qy 286 CTGATGATGTTCTTGTGGGTATCCCGCATGTTCGGTGTTCGTAATCTATCTGATG 345

Db 247 CTGATGATGTTCTTGTGGGTATCCCGCATGTTCGGTGTTCGTAATCTATCTGATG 306

Qy 346 CCGCTGCAAAATCGCGGCTCCGGATATGCGCTTCCCGGATGAAACAACCTGTCGTTCTG 405

ALIGNMENTS

RESULT 1

```

US-10-127-032-32
; Sequence 32, Application US/10127032
; Publication No. US20030113742A1
; GENERAL INFORMATION:
; APPLICANT: Whiteley, Marvin
; APPLICANT: Bangera, M. Gita
; APPLICANT: Lory, Stephen
; APPLICANT: Greenberg, Everett Peter
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF
; FILE OF INVENTION: BIOFILM FORMATION
; FILE REFERENCE: UIZ-070CP
; CURRENT APPLICATION NUMBER: US/10/127,032
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/285,190
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: US 60/344,142
; PRIOR FILING DATE: 2001-10-24
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 1593
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-10-127-032-32

```

```

Sequence 508, App
Sequence 524, App
Sequence 529, App
Sequence 529, App
Sequence 18, Appl
Sequence 31, Appl
Sequence 101, App
Sequence 133, App
Sequence 144, App
Sequence 147, App
Sequence 153, App
Sequence 153, App
Sequence 240, App
Sequence 285, App
Sequence 286, App
Sequence 454, App
Sequence 454, App
Sequence 62, Appl
Sequence 94, Appl
Sequence 205, App
Sequence 289, App
Sequence 341, App
Sequence 54, Appl
Sequence 287, App
Sequence 452, App
Sequence 452, App
Sequence 463, App
Sequence 59, Appl
Sequence 430, App
Sequence 266288,
Sequence 266288,
Sequence 266288,
Sequence 382, App
Sequence 269, App

```

```

13 398.8 23.8 16569 16 US-10-308-264-508
14 398.8 23.8 16569 16 US-10-308-264-524
15 398.8 23.8 16569 16 US-10-308-264-529
16 398.8 23.8 16570 16 US-10-308-264-18
17 398.8 23.8 16570 16 US-10-308-264-31
18 398.8 23.8 16570 16 US-10-308-264-101
19 398.8 23.8 16570 16 US-10-308-264-133
20 398.8 23.8 16570 16 US-10-308-264-144
21 398.8 23.8 16570 16 US-10-308-264-147
22 398.8 23.8 16570 16 US-10-308-264-153
23 398.8 23.8 16570 16 US-10-308-264-240
24 398.8 23.8 16570 16 US-10-308-264-285
25 398.8 23.8 16570 16 US-10-308-264-286
26 398.8 23.8 16570 16 US-10-308-264-454
27 398.8 23.8 16571 16 US-10-308-264-26
28 398.8 23.8 16571 16 US-10-308-264-62
29 398.8 23.8 16571 16 US-10-308-264-94
30 398.8 23.8 16571 16 US-10-308-264-205
31 398.8 23.8 16571 16 US-10-308-264-289
32 398.8 23.8 16571 16 US-10-308-264-341
33 398.8 23.8 16572 16 US-10-308-264-54
34 398.8 23.8 16572 16 US-10-308-264-287
35 398.8 23.8 16572 16 US-10-308-264-535
36 398.8 23.8 16573 16 US-10-308-264-452
37 398.8 23.8 16573 16 US-10-308-264-463
38 398.8 23.8 16574 16 US-10-308-264-59
39 398.8 23.8 16574 16 US-10-308-264-430
40 398.4 23.8 2174 13 US-10-027-632-266288
41 398.4 23.8 2174 13 US-10-027-632-266289
42 398.4 23.8 2174 15 US-10-027-632-266288
43 398.4 23.8 2174 15 US-10-027-632-266289
44 398.4 23.8 16568 16 US-10-308-264-382
45 398.4 23.8 16572 16 US-10-308-264-269

```


Db 655 GTGCTGCGCGGCTGGTGAACATGATGCTGATGAGATCCACTTCGCGACCAAGCTTCTTC 714

Qy 766 AATCCTCTGCGCGGCTGAGCCGATTTCTGTACCAACATCTCTGTTGGTTCTTTGGGCAC 825

Db 715 AGCGCGCGCGGCGGCGACCCGGTGTGTTCAGCACGTTGTCTGTTCTTCTGCGGCAC 774

Qy 826 CCGAAGTGATACATCATCTCGCGGCTTTGGCATCATCAGCATGCTGCTGCGAC 885

Db 775 CCCAGGTGTACATCATGATCTCGCGGCTTTGGCATCATCAGCATGCTGCTGCGAC 834

Qy 886 TTCTCGAAAAAGCGGCTTCTCGGTATACCTGCGGATGTTGCAATGTTGGCAATCGGT 945

Db 835 TTGCGCGCAAGCGCTGTTCTCGGTACACCTCGATGTTGACCGACCGCGCATCGCC 894

Qy 946 GTTCTGGGCTTTGCTGCTGGCGGCAACATGATACACCGTTGTATGTGCTGACCCAG 1005

Db 895 TTCTCTCTCTTCTGTTGCTGGCGCACCATGTTCTGTTGTCGTCATCCCGGTACCGGC 954

Qy 1006 CAATCTCTACTTCTGCTGCGGCAACATGTTGATCGCGGTGCGGACCGGATTAAGATCTTC 1065

Db 955 GAGCTGTTCTTCTGATGAGCCACCATGCTGATCGCGGTGCGGACCGGTTGAGGTGTC 1014

Qy 1066 TCGTGATGCGCAGATGTTGGGCGGCTCGGTTGAGTTCAATCGCGATGCTCTGGGCC 1125

Db 1015 AACTGGGTGACCAACATGTTGGAGGTTTCGCTGACCTTCGAGACGCGCATCTGTTGCGC 1074

Qy 1126 TTGCTTTATGTTCTGTTTACCGTGGTGTGACCGGTATGCTGCTGCGGCAAGCG 1185

Db 1075 GTGGCTTCTGCTATCTCTGTTTACCATCGCGGCTTCTCGGACTGATGCTGGCATCGCC 1134

Qy 1186 GGTCTGACCGTGCATATCAGACACCTATTACGTTGGTGGGCACTTCCATTTATGTATG 1245

Db 1135 CGGCGAGACTTCCAGTACCAAGACACTTACTCGTGGTTCGCGCACTTCCACTAGTGTG 1194

Qy 1246 TCGCTGGTGGATCTTTGGCATCTTTCGCGGTATCTACTTCTATACATGCGCAAGTTCTCG 1305

Db 1195 GTGCGCGCGCGATCTCGGCACTTTCGCTCGGCTTACTACTGCTGCCGAGTGGAC 1254

Qy 1306 GCGCGGCTTTTCCGGAATGGGCTCAAAAGTGCATTTCTGACCTTCTTCACTCGGTGCG 1365

Db 1255 GGCACATAGACAGACACCTCGGCAAGCTGCACTTCTGATGAGCTTCACTGGCATG 1314

Qy 1366 AACGTACGTTTCTCGCGAGCACTTCTGGGAGCTGAGGTATGCGCGGCTTATCATC 1425

Db 1315 AACCTGGGCTTCTCCCATGCACTTCTGCGGCTTCGCGGATCGCGGATCGCGGATCGCC 1374

Qy 1426 GACTATCCGAGAGCTTCGCGCTGTGGAACAAAGTCTGCTCTATGCTGCTGCTGCGC 1485

Db 1375 GACTAACCTGAGTTCCCGCACTTCAACATGTTCTGCTGATCGGCGCTTCATG--- 1431

Qy 1486 TTGCGCTCGTCTCTGTTCTTCACTGATGATCTTTGTTGTTATAGCTGTTGCTGGCGCGC 1545

Db 1432 TTGCGCACCAACAGCTGCTGTTCTGTTTCACTGCTATCAAGTGCATCCGCGGCGCAAG 1491

Qy 1546 GAGACCGTTCGAACCGTGGGCGAATTCGCGGATACGCTGGATGGACGCTGCCATCA 1605

Db 1492 CCGGCGGCTTCCAAAGCCCTGGGAGC---CGCGGAGGCGCTTGGAGTGAGCATCCCTCG 1548

Qy 1606 CGGCTTCGCGGCGCACACGTTTCAAAACGCTGCGCAAG 1641

Db 1549 CCGGCGGCTTACACACCTTCAGACCCCGCGCGAG 1584

RESULT 3

US-10-282-122A-25546

; Sequence 25546, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Karl

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA 034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 25546

; LENGTH: 1629

; TYPE: DNA

; ORGANISM: Mycobacterium avium

US-10-282-122A-25546

Query Match 24.3%; Score 406.2; DB 16; Length 1629;

Best Local Similarity 59.0%; Pred. No. 1.3e-112;

Matches 800; Conservative 0; Mismatches 533; Indels 24; Gaps 5;

Qy 288 GATGATGTTCTTTGTGGGTATCCCGATTCGCGGTTTGGTGGTAACTATCTGATGCC 347

Db 150 GATCATCTCTCTGCTGTACCGCACCCCGGTGTTTCGGTTTCGCCAACCTGGTGTGCC 209

Qy 348 GCTCAATCGCGCTCCGATATGGCTTCCCGCGTATGAACAACTGCTGTTGGCT 407

Db 210 GCTGAGATCGCGGCGCGACGTTGCGCTTCCCGCGCTCAACGCTTCTGTTCTGGCT 269

Qy 408 GTTCATTGCGGTACCGCATGGGCGTGGCTTCTGTTGCGACCGGCGGTGACGCTCA 467

Db 270 GTTCCTGTTGCGGCGGTGATCGCGGCTCCGCTTTCATCTGCGGCGGCGCGCCGA 329

Qy 468 GCTGGGTTGCGGCTTGTGGTTCTGTACCGCGCTGTCGACCCCGGAAGTGTGCTA 527

Db 330 CTTCGGTGTGACCGCTTACACCCCGCTGTGTGACCGCTGCTGCTGATCATGGCGC 386

Qy 528 TTGATGGAACCTCGGATTTTCGCGGTTCACTTGTTCGGGTGCTTCTGCTGATCATGGCGC 587

Db 387 ---GGGAGACCTGTGGATCACCGGCTGTGTCGCGGCTGCGGACCATCTCTGGGTGC 443

Qy 588 GATCAACATGATCAGACCTTCTTGAACATCGCGGCGCGCGGATGACGCTGCAAAAAGT 647

Db 444 GGTCAACATGATCAGACCTGTTCTGATTCGCGCGCGCGCGGATGACGATGTTCCGGAT 503

Qy 648 GCCGTGTTCTGCTGGTGCATCTTTATCAGGCTTGGCTGATCTGCTGCTGCTGCTGCTGCT 707

Db 504 GCCGATCTTCACTTGGAAATCTCTGGTGACCTCGATCTGATCTGATCTGATCTGCTTCCGAT 563

Qy 708 TCTGGCTGTGCAATCACCATTGCTGTCGACCGACCGTAACCTTTCGCGACGACCTTCTTCAA 767

```
Db 564 CCTGACCGCGCGTGTTCGGGCTGGCGCGCCGACCGGCATCTGGCGCCCATGCTACGA 623
Qy 768 TCCTGCTGGCGCGGTGACCGATTCCTGTACCAACACATCTGTGGTCTTTTGGGACCC 827
Db 624 CGCGCCCAACGGCGAGTTC---TGTGTGGCAGCACCTGTTCTGGTCTTCGGCCATCC 680
Qy 828 GGAAGTGTACATCATCAATTCGCCGCTTTGGCGATCATGAGCATGCTGTGTCGACCTT 887
Db 681 CGAGGTGTACATCATCGGTTCGGTCTTCGGGATCATCAACGAGATCATCCGCTGT 740
Qy 888 CTCGAAAAGCGGCTTCCTGCTACCTGCGATGGTCTAGCAATGGTGCAATCGTGT 947
Db 741 CGCGCGAAGCGGCTGTTCGGCTACACACGCTGGTGTAGCGACGCTGTCGATCGCGC 800
Qy 948 TCTGGGCTTTGCTGCTGGCGGACCAACATGTATACACGCTTGTGTATGTCGCTGACCCAGCA 1007
Db 801 GCTCTCGGTGCGGTGTGGCGGACCAACATGTTGCGCACCGAGCGCTTCTGCTCGCTT 860
Qy 1008 ATCCTACTTATCTGCGCCACCATGATGATCGCGGTGCGGACCGGCATTAAGATCTTCTC 1067
Db 861 CTTTCTGTTTATG-----ACGTACCTGATCGCGGTGCGGACCGGATCAAGTCTTCAA 914
Qy 1068 GTGGATCGCACGATGTGGCGGCTCGGTGTAGTTCAAAATCGCGATGCTCTGGSCCTT 1127
Db 915 CTGGATCGGCACGATGTGGAAGGGGAGTTGACATTTGAGACCGCATGCTGTCTGCGT 974
Qy 1128 TGGCTTTATGTTCTGTTACCTGCGGTGTGTGACCGGTATCGTCTGGCCCAAGCGG 1187
Db 975 CGGCTTCTGCTCACTTCTGCTGCTGCTGTGCTGTGACCGCGGTATGCTGCGCCAGCCGC 1034
Qy 1188 TCTGGACGCTGATATCAACACACCTATTACGTTGTGGCGCTTCCATTTATGATGTC 1247
Db 1035 GCTGGACTTCAAGTACCGACACCTATTTCGTTGGTGGCGACTTCCACTAGCTGCTGT 1094
Qy 1248 GCTGGGTGCGATCTTTGCGATCTTCGCGGTATCTACTTTTATACGCGGAAGTTCCTCGG 1307
Db 1095 CGGCACCATCGTGTTCGCCACCTTCGGGGGGTCTACTTCTGTTCCCGAAGTACCGG 1154
Qy 1308 CCGCGCTTTCGCGAATGGGCTCAAGCTGCACTTCTGACCTTCTTATCGGTGCGNA 1367
Db 1155 CCGCTGCTCGACGCGGTGGGCAAGCTGCACTTCTGTTGAGCTTCTATCGGTTCCA 1214
Qy 1368 CGTCAGCTTCTTCCCGACGACTTCTTGGGACGTCAGGCTATGCGCGCGCTTATCATGA 1427
Db 1215 CACACCTTCTGCTGACGACTGCTGGCGGATTTGGGCATGCGCGCGCTACGCGGA 1274
Qy 1428 CTATC-----CCGAGCCTTCGCGTGTGGAAACAAAGTCTGCTATGTTGGTCTCT 1481
Db 1275 CTACCTGCGCAGCGGCTTCCAGCCCTACACGTCGCTCCACGCTGCGGCGCTTCAT 1334
Qy 1482 GGCCTTCGCTCGCTCTGCTTCTTCTTCTGCTGATCTTGTCTATACGCTGTTGCTGGCG 1541
Db 1335 CTTGGGCGCTCGATGTTCCGTTCTGCTTGGAACTCTTCAAGAGCTGCGCTACGCGGA 1394
Qy 1542 CCGCGAGACCGTTCGAAACCGTGGGCGAATTCGCCGATACGCTGGAATGACGCTGCC 1601
Db 1395 GGTGCTGACCGTGGAGACCGCGTGGG---TTACGGCAACTGCTGGAGTGGCGACAG 1451
Qy 1602 ATCACCCTTCGCGGCCACACGTTTCGAAACGCTGCC 1638
Db 1452 CTGCGCGCGCGCGGACCAACTTCCAGCTGCC 1488
```

RESULT 4

```
US-10-308-264-150
; Sequence 150, Application US/10308264
; Publication No. US20040029133A1
; GENERAL INFORMATION:
; APPLICANT: Herinstad, Corinna
; TITLE OF INVENTION: MITOCHONDRIAL DNA POLYMORPHISM
; FILE REFERENCE: 660088.461
; CURRENT APPLICATION NUMBER: US/10/308,264
```

```
; CURRENT FILING DATE: 2002-11-25
; NUMBER OF SEQ ID NOS: 697
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 150
; LENGTH: 16559
; <211> conflict
; TYPE: DNA
; ORGANISM: Homo sapiens mitochondrial
; FEATURE:
; LOCATION: 3106
; OTHER INFORMATION: "n" refers to a nucleotide at position 3107 of the
; OTHER INFORMATION: Cambridge Reference Sequence (CRS, Anderson et al.
; OTHER INFORMATION: 1981 Nature 290:457-465), later determined to be
; OTHER INFORMATION: absent from the human mtDNA sequence according to
; OTHER INFORMATION: the revised CRS (+CRS, Andrews et al. 1999 Nature
; OTHER INFORMATION: Genetics 23:147); "n" therefore represents a
; OTHER INFORMATION: single nucleotide deletion at a position corresponding to
; OTHER INFORMATION: nucleotide 3107 of SEQ ID NO:1 as provided herein.
US-10-308-264-150
```

```
Query Match 23.8%; Score 398.8; DB 16; Length 16559;
Best Local Similarity 58.2%; Pred. No. 5e-110;
Matches 744; Conservative 0; Mismatches 522; Indels 12; Gaps 2;
```

```
Qy 244 AACGACACCTGTGGAACGTCATGTTACCTACCATGATCTCTGATGATGTTCTTTGTG 303
Db 6050 AACGACACCATCTACAACGTTATCGTCACAGCCCATGATTTGTAATAATCTTCTTCATA 6109
Qy 304 GGTATCCCGCATTTTCGGTGGTTTTGTAACATCTGATGCGCGTCAAAATCGGCGCT 363
Db 6110 GTAATACCATCATATTCGGAGGCTTTGGCACTGACTAGTTCCTTAATAATCGGTGCC 6169
Qy 364 CCGGATATGGCTTCCCGCGTATGAACAACTGTCGTCTGGCTGTTTCATTGCGGTACC 423
Db 6170 CCGGATATGGCTTCCCGCGTAAACAACTGACTTCTGACTTACCTCCCTCTCTC 6229
Qy 424 CGGATGGCGTGGCTTCGCTTTCGACCGGCGGTGACGTCAGCTGGGTTCGGGCGTT 483
Db 6230 CTACTCTGCTCGCATCTGCTATAGTAGAGCGCGGACAGAAACAGTTGAAACAGTC--- 6286
Qy 484 GGTGGGTTCGTACCGCGCTGTCGACCGGGAAGCTGGCTATTGATGAGACCTCGCG 543
Db 6287 -----TACCTCCCTTAGCAGGAACTACTCCACCCCTGGAGCTCCGTAGACTAAC 6340
Qy 544 ATTTTCGGGTTCACTTGTCGGGTGCTCTCGATCATGGCGCGATCAACATGATCAG 603
Db 6341 ATCTTCTCTTACACCTAGCAGGTGCTCTCTATCTTAGGGGCCATCAATTTTCATACA 6400
Qy 604 ACCTTCTTGAACATGCGCGCCCGGCATGACGTCGACAAAGTGGCGTGTCTCGTGG 663
Db 6401 ACAATATCAATATAAAACCCCTGCCATAACCAATACCAACGCCCTCTTCGTCGA 6460
Qy 664 TCGATCTTATACAGGTTCGCTGATCTGCTGGCGTGGCGTCTGGCTGTGCAATC 723
Db 6461 TCGGTCCTAATACAGCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 6520
Qy 724 ACATGCTGTCGACCGACCTTAACCTTCGGCACCACTTCTTCAATCTGTCGGCGGCT 783
Db 6521 ACTATACTACTACAGACCGCAACCTCAACACCACTTCTTTCAGCCCGCGGAGGGA 6580
Qy 784 GACCGGATCTGTACCAACATCTCTGTGTTCTTTGGGCACCGGGAAGTGTACATC 843
Db 6581 GACCCCATCTATACCAACACCTATTCTGATTTTTCGGTCACTCTCTCTCTCTCTCTCT 6640
Qy 844 ATTTCTCGCGCTTTTGGCATCATCAGCATGTCGTCGACCTTCTC---GAAAGACCG 900
Db 6641 ATCTTACAGGCTTCGGAATATCTCCATATTTGTAACCTTACTTCCCGAATAAAGAA 6700
Qy 901 GTCTCGGTTACTGCGGATGTTCTATGCAATGGTGGCAATCGGTGTTCTGGGCTTTGTC 960
Db 6701 CCATTTGGATACATAGTATGTTCTGAGCTATGATCAATTAATTAATTAATTAATTAAT 6760
Qy 961 GTCTGGGCGCACCATGTACACCGTGGTATGTCGCTGACCCGCAATCTTACTTCTATG 1020
```


Db 6761 GTGTGAGCACACCATATATTTACAGTAGGAATAGACGTAGACACACGAGCATATTTCCACC 6820
Qy 1021 CTGGCCACCATGGTGTGATCGGGTGGCGACCGGATTAAGATCTTCCTGGATCGCCAG 1080
Db 6821 TCCGCTACCATTAATCATCGTATCCCGACCGGGGTCAAGTATTTAGCTGACCTGCCACA 6880
Qy 1081 ATGTGGGGGGCTCGGTGTAGTTCAAAATCCCGGATGCTCTGGGCCCTTTGGCTTTATGTT 1140
Db 6881 CTCCACGGAAGCAATATGAATGATCTGTGCTGATGCTCTGAGCCCTAGGATTCATCTTT 6940
Qy 1141 CTGTTTACCGTGGGTGTGACCGGTATCGTGTGCTGCGCCCAAGCGGTCTGGACCGTGCA 1200
Db 6941 CTTTTTACCGTAGTGGCTGACTGGATTTGATTAGCAAACTCATCATAGACATCGTA 7000
Qy 1201 TATCAGACACCTATTACGTGGTGGGCGCATCTTCATATATGATGCTGGCTGGGTGCGATC 1260
Db 7001 CTACACGACACGTACTAGTTAGCTTCACTTCCATCATATTCATCGGCGTAAATCTAACTTCTTC 7060
Qy 1261 TTTCCGATCTTCGGCGGTATCTACTTTTACATCGGAAGTTCTCGGGCGCGCTTTCCCG 1320
Db 7061 TTTGCCATCATAGAGGCTTCATCTACTGATTTCCCGTATTTCTAGGCTTACACCTAGAC 7120
Qy 1321 GAATGGGCTGAAAGCTGCTCTGAGCCTTTCTGAGCCTTTCTCATGGTGGCAACGTCTCTTC 1380
Db 7121 CAAACCTACGCCAAATCCATTTCCGTATCATATTCATCGGCGTAAATCTAACTTCTTC 7180
Qy 1381 CCGGACGACTTCTCGGAGCTCAGGCTATCGCGCGCTTACATCGACTATCCCGAAGCC 1440
Db 7181 CCACACACATTTCTCGGCTTCTCGGATTCGGATATCGGATGCCCCGAGCTTACTCGGACTACCCGATGA 7240
Qy 1441 TTCCGCTGTGGAAACAAAGTCTCGTCTATGTTGGTGGTCTCGGCTTCGCTCTGTTCTG 1500
Db 7241 TACACACATGAATAATCTCTATCATCTGTAGGCTCATTCATTTCTTAACAGCAGTAATA 7300
Qy 1501 TTCTTTCATCGTGAATCTTT 1518
Db 7301 TTAATAATTTTCATGAT 7318

RESULT 5

US-10-308-264-145
; Sequence 145, Application US/10308264
; Publication No. US20040029133A1
; GENERAL INFORMATION:
; APPLICANT: Herinstdt, Corinna
; TITLE OF INVENTION: MITOCHONDRIAL DNA POLYMORPHISM
; FILE REFERENCE: 660088.461
; CURRENT APPLICATION NUMBER: US/10/308,264
; CURRENT FILING DATE: 2002-11-25
; NUMBER OF SEQ ID NOS: 697
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 145
; LENGTH: 16561
; <211> conflict
; TYPE: DNA
; ORGANISM: Homo sapiens mitochondrial
; FEATURE:
; LOCATION: 3108
; OTHER INFORMATION: "n" refers to a nucleotide at position 3107 of the
; OTHER INFORMATION: Cambridge Reference Sequence (CRS, Anderson et al.
; OTHER INFORMATION: 1981 Nature 230:457-465), later determined to be
; OTHER INFORMATION: absent from the human mtDNA sequence according to
; OTHER INFORMATION: the revised CRS (rCRS, Andrews et al. 1999 Nature
; OTHER INFORMATION: Genetics 23:147); "n" therefore represents a
; OTHER INFORMATION: single nucleotide deletion at a position corresponding to
; OTHER INFORMATION: nucleotide 3107 of SEQ ID NO:1 as provided herein.
US-10-308-264-145

Query Match 23.8%; Score 398.8; DB 16; Length 16561;
Best Local Similarity 58.2%; Pred. No. 5e-110;
Matches 744; Conservative 0; Mismatches 522; Indels 12; Gaps 2;

Qy 244 AACGGACACCTGTGGACAGTCATGTTTACCTACATGGTATTTCTGATGATGTTCTTTG 303
Db 6052 AAGGACACATCTACAAAGTATTCGTACACAGCCATGATTTGTAATAATCTTTTCA 6111
Qy 304 GGTATCCCGCATGTTGGTGGTGGTAACTATCTGATGCCGTGCAATGCGGCT 363
Db 6112 GTAAATCCCATATAATCGGAGGCTTTGGCACTGACTAGTTCCCTTAATAATCGGTGCC 6171
Qy 364 CCGGATATGCGCTTCCCGGTATGAACAACCTGTCTGTTCTGGCTGTTCATTTCCCGGTACC 423
Db 6172 CCGGATATGCGTTCCTCCGCAATAAACAATGAAGCTTCTGACTCTTACCTCCCTCTCTC 6231
Qy 424 GCGATGGGCTGGCTGGCTGTTCGACCGGGGGTGAAGTTCAGTGGGTGGGCGCTT 483
Db 6232 CTACTCTGCTGCACTCTGCTATAGTGGAGGCGGAGAGGTAACAGTTGAACAGTCT 6288
Qy 484 GGTGGGTTCTGTACCGCGCTGTGACCGCGAGCTGGCTATTCGATGGACCTTCGGG 543
Db 6289 -----TACCCTCCCTTAGCAGGAACTACTCCACCCTGGAGCTCCGTAGACCTAAC 6342
Qy 544 ATTTTCGGGTTTCACTGTGGGTGGCTCTGCTGATCATGGGCGCATCAACATGATCAG 603
Db 6343 ATCTTCTCTTACACCTAGCAGTGTCTCTCTATCTTAGGGGCGCATCAACTCATAC 6402
Qy 604 ACCTTCTTGAACATGCGCGCGCGCGCATGAGCTGCACAAAGTGGCGTTGTTCTCGTGG 663
Db 6403 ACAATTATCAATATAAAACCCCTGCCATAAACCATAACAAACGCCCTCTTCGCTGA 6462
Qy 664 TCGATCTTTATCAGGCTTGGCTGATCTCTGGGCGCTCCCGGTTCTGGCTGGTCAATC 723
Db 6463 TCGCTCTTAATACAGCAGTCTTCTCTCTATCTCTCCAGTCTAGCTCTGGCATC 6522
Qy 724 ACCATGCTGTGACCGACCGCTTAATCTTGGGACAGCCTTCTCAATCTCTCTGGGCGGCT 783
Db 6523 ACTATACTACTAACAGACCGCACTCAACACACCTTCTTCGACCCCGCGGAGGGA 6582
Qy 784 GACCGATTTCTTACCAACACATCTCTGTGTTCTTTGGGACCCCGGAAGTGTACATCATC 843
Db 6583 GACCCATTTCTTACCAACACCTTATCTGATTTTTCGGTCAACCCCTGAGTTTATATCT 6642
Qy 844 ATTTGCGCGGCTTTGGGCTATCATGACCATGTGCTGCTGACCTTCTC-----GAAAAAGCG 900
Db 6643 ATCTACCGGCTTCCGAATATCTCCATATTTGTAACCTTACTCTCCGGAATAAAGAA 6702
Qy 901 GTCTTGGTTTACCTGGGATGTTCTATGCAATGGTGGGAACTCGGTGTTCTGGGCTTTGTC 960
Db 6703 CCAATTTGGATACATAGTATGCTGAGCTATGATATCAATTTGGGCTTCTAGGGTTTATC 6762
Qy 961 GTCTGGGCGCACCATGTACACCGTTGTTGTTGCTGACCCAGCAATCTTACTTCTATG 1020
Db 6763 GTGTGAGCACACCATATATTTACAGTAGGAATAGAGTAGACACAGAGCATATTTTACC 6822
Qy 1021 CTGGCCACCATGTGTATGCGGGTGGCGGAGCTTGAATCTTCTGTTGGATCGGACG 1080
Db 6823 TCCGCTTACCAATAATCATCTCTATCCCGCGGCTCAAAAGTATTTAGCTGACTCGCCACA 6882
Qy 1081 ATGTGGGGGCGCTCGGTTGAGTTCAAAATCGCGGATGCTCTGGGCGCTTTGGCTTTATGTT 1140
Db 6883 CTCACGGAAGCAATATGAATGATCTGCTGAGTGTCTGAGCCCTAGGATTCATCTT 6942
Qy 1141 CTGTTTACCGTGGTGTGTGACCGTATCTGCTGGCCCAAGCGGCTTGAACCGTGA 1200
Db 6943 CTTTTCACCGTAGTGGCTGACTGCGCATTTGATTAGCAAACTCATCACTAGACATCGTA 7002
Qy 1201 TATCAGACACCTTATTTACGTGGTGGCGCACTTCCATTTATGATGCTGGTGGGTCGATC 1260
Db 7003 CTACACGACAGTACTACCTGTTAGTCTCACTTCCACTATGCTCTATCAATAGAGCTGTA 7062
Qy 1261 TTTGCGATCTTCCCGGTATCTACTTTTACATGCGGAAGTCTCGGCGCGCGCTTTCCCG 1320
Db 7063 TTTGCGATCATAGAGGCTTCACTTCACTGATTTCCCGCTATTTCTAGGCTACACCTAGAC 7122
Qy 1321 GAATGGGCTGCAAAAGCTGCATCTTCTGAGCTTCTTCTATCGGTGCGAAGCTCACGTTCTTC 1380

Db 7123 CAACCTACGCGCAAAATCCATTTCACTATCATATATCATCGCGGTAAATCTAACTTTCTTC 7182
 QY 1381 CGCAGCACTTCTCGGGACGTCAGGATATCGCGGCGGTATACATCGACTATCCCGAAGCC 1440
 Db 7183 CCACAACACTTCTCGGCTATCCGGAATGCCGAGCTTACTCGGACTACCCCGATGCA 7242
 QY 1441 TTCGGCTGTGGACAAAGTCTCGTCTATGGTGGGTTCTGGCTTCGGCTTCGCTG 1500
 Db 7243 TACACCACATGAACATCCTATCATCTGTAGGCTCAATTCATTTCTTAACAGCAGTAATA 7302
 QY 1501 TTCCTCATCGTGATCTTT 1518
 Db 7303 TTAATATTTTCATGATT 7320

RESULT 6

US-10-308-264-414
 ; Sequence 414, Application US/10308264
 ; Publication No. US20040029133A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Herinstdt, Corinna
 ; TITLE OF INVENTION: MITOCHONDRIAL DNA POLYMORPHISM
 ; FILE REFERENCE: 660088.461
 ; CURRENT APPLICATION NUMBER: US/10/308,264
 ; CURRENT FILING DATE: 2002-11-25
 ; NUMBER OF SEQ ID NOS: 697
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 414
 ; LENGTH: 16568
 ; <211> conflict
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens mitochondrial
 ; FEATURE:
 ; LOCATION: 3106
 ; OTHER INFORMATION: "n" refers to a nucleotide at position 3107 of the
 ; Cambridge Reference Sequence (CRS, Anderson et al.
 ; OTHER INFORMATION: 1981 Nature 290:457-465), later determined to be
 ; OTHER INFORMATION: absent from the human mtDNA sequence according to
 ; OTHER INFORMATION: the revised CRS (rCRS, Andrews et al. 1999 Nature
 ; OTHER INFORMATION: Genetics 23:147); "n" therefore represents a
 ; OTHER INFORMATION: single nucleotide deletion at a position corresponding to
 ; OTHER INFORMATION: nucleotide 3107 of Seq ID No.1 as provided herein.
 US-10-308-264-414

Query Match 23.8%; Score 398.8; DB 16; Length 16568;
 Best Local Similarity 58.3%; Pred. No. 5e-110;
 Matches 744; Conservative 0; Mismatches 522; Indels 12; Gaps 2;
 QY 244 AAGGACACCTGGAGCTGATGTTACCTACCATGGTATCTGATGATGTTCTTTGG 303
 Db 6050 AAGGACACATCTACAACTGATCGTCAGAGCCATGCTATTAATCTTCTTATA 6109
 QY 304 GGATCCCCCATGTTTCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 363
 Db 6110 GATATACCATCATATATCGAGCTTGGCACTGACTAATTCCTTAAATATCTTCTATA 6169
 QY 364 CCGGATATGCGCTTCGCGGTATGAACAACTGTGTTCTGGCTGTGTTCTGCGGTACC 423
 Db 6170 CCGGATATGCGCTTCGCGGTATGAACAACTGTGTTCTGGCTGTGTTCTGCGGTACC 6229
 QY 424 GCGATGGGCTGCTGCGGTATGCGACCGGGGTGAGCGTACGCTGGGTTCGGGGCTT 483
 Db 6230 CTACTCCTGCTGCTATGCTATGAGGAGCGGAGAGGAGGAGGAGGAGGAGGAGGAG 6286
 QY 484 GGTGGGTTCTGTACCCGCGCTGTGCGACCGCGGAGCTGGCTATTCATGAGCACTCGCG 543
 Db 6287 -----TACCCCTCCCTTAGCAGGAACTACTCCACCCCTGGAGCCTCCGTAGACCTAAC 6340
 QY 544 ATTTTCGGGTTCACTTGTGGGTGCTTCTGATCATGGGGCGCATCAACATGATACG 603
 Db 6341 ATCTTCTCTTACACCTAGCAGGTGCTCTCTATCTTAGGGGCGCATCAATTTTCATACA 6400

QY 604 ACCTTCTTGAACATGCGCGCCCGCATGACCTGCAAAAGTCCCGTGTGTTCTCGTGG 663
 Db 6401 ACAATATCAATATAAAACCCCTCGCATACCAATACCAAAACCCCTCTTCTGCTGA 6460
 QY 664 TGAATCTTTATCAGCGCTGGCTGATCTCTGCTGGCGCTGCGGTTCTGGCTGGTGAATC 723
 Db 6461 TCGGCTCTAATCAGACGAGTCTACTTCTCTCTATCTCTCCAGTCTTACTGCTGGCATC 6520
 QY 724 ACCATGCTGCTCAGCAGCGTAATCTTGGCAGACCTTCTCAATCTCTGCTGGCGGCT 783
 Db 6521 ACTATACTACTAACAGACCGCAACCTCAACACCTTCTTCGACCCCGCGGAGGAGA 6580
 QY 784 GACCGATTTCTGTACCAACATCTCTGTTCTTTGGGCACCCGGAAGTGTACATCATC 843
 Db 6581 GACCCCATTTCTATACCAACACCTATCTGATTTTCGGTCACTGAAAGTTTATATCTT 6640
 QY 844 ATTTCTGCGCGCTTTGGCATCATCAGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 Db 6641 ATCTTACAGGCTTCGGAATATCTCCCATATTTGAACTTACTCTCGGAAATAAGAA 6700
 QY 901 GTCTTCGGTTACTCGCCGATGCTCTATGCAATGGTGGCAATCGTGTCTGGGCTTTGTC 960
 Db 6701 CCATTTGGATACATAGGTATGTTCTGAGCTATGATATCAATTTGCTTCTAGGTTTATC 6760
 QY 961 GTCTGGGCGCACCATGATACACGTTGGTATGCTGCTGACCCAGCAATCTCTACTCATG 1020
 Db 6761 GTGTGAGCACCATATATTTTACAGTAGGAATAGACGTAGACACACGAGCATATTTTACC 6820
 QY 1021 CTGGCCACCATGTTGATCGCGTCCGACCGGCAATTAAGATCTTCTGCTGATCGCCACG 1080
 Db 6821 TCCGCTACCATATCATCTGCTATCCCAACCGGCTCAAGTATTTAGTACTGACTCGCCACA 6880
 QY 1081 ATGTGGGCGGCTCGTGTGATTTCAATTCGCCGATGCTCTGGGCTTTGGCTTTTATGTT 1140
 Db 6881 CTCCAGCGAAGCAATATGAAATGATCTGCTGAGTCTGCTGAGCCCTAGGATTCATCTTT 6940
 QY 1141 CTGTTACCGTGGTGGTGGACCGGTATGCTGCTGCCCAAGCGGCTCTGGACCGTGA 1200
 Db 6941 CTGTTACCGTGGTGGTGGACCGGTATGCTGCTGCCCAAGCGGCTCTGGACCGTGA 7000
 QY 1201 TATCAGCACCTATTAAGTGGTGGCGCACTTCATTTATGATGCTGCTGGTGGCGATC 1260
 Db 7001 CTACACGACGTAAGTGGTGGCGCACTTCATTTATGATGCTGCTGGTGGCGATC 7060
 QY 1261 TTTGCGATCTCGCGGTATCTACTTTTACATCGCGAAGTCTTGGCGCGGCTTTCCCG 1320
 Db 7061 TTTGCGATCTAGAGGCTTCAATCTGATTTTCCCTTCTGAGGCTACACCTAGAC 7120
 QY 1321 GAATGGGCTGCAAAAGTGCATCTTCTGACCTTCTTCACTCGTGGCAACGTCAGCTTCTTC 1380
 Db 7121 CAACCTTACGCCAAATTCATTTCACTATCATATTCATCGCGTAAATCTAACTTTCTTC 7180
 QY 1381 CCGGACGACCTTCTGGGACGTCAGGATGTCGGCGGCTTACATCGACTATTCGCGAGCC 1440
 Db 7181 CCACACACTTTCTCGGCTTATCCGGAATGCCCGGAGCTTACTCGGACTACCCGATGA 7240
 QY 1441 TTTGCGCTGTGGAAACAAAGTCTGCTCTATGCTGGTTCCTGCGCTTGGCTTCTGCTGCTG 1500
 Db 7241 TACACCATGAAACATCTCTATCATCTGAGGCTCATTCATTTCTCTAACAGCAGTAATA 7300
 QY 1501 TTCTTCACTGATCTTT 1518
 Db 7301 TTAATATTTTCATGATT 7318

RESULT 7

US-10-308-264-181
 ; Sequence 181, Application US/10308264
 ; Publication No. US20040029133A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Herinstdt, Corinna
 ; TITLE OF INVENTION: MITOCHONDRIAL DNA POLYMORPHISM
 ; FILE REFERENCE: 660088.461

```

; CURRENT APPLICATION NUMBER: US/10/308,264
; CURRENT FILING DATE: 2002-11-25
; NUMBER OF SEQ ID NOS: 697
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 181
; LENGTH: 16569
; <211> conflict
; TYPE: DNA
; ORGANISM: Homo sapiens mitochondrial
; FEATURE:
; LOCATION: 3107
; OTHER INFORMATION: "n" refers to a nucleotide at position 3107 of the
; OTHER INFORMATION: Cambridge Reference Sequence (CRS, Anderson et al.
; OTHER INFORMATION: 1981 Nature 290:457-465), later determined to be
; OTHER INFORMATION: absent from the human mtDNA sequence according to
; OTHER INFORMATION: the revised CRS (rCRS, Andrews et al. 1999 Nature
; OTHER INFORMATION: Genetics 23:147); "n" therefore represents a
; OTHER INFORMATION: single nucleotide deletion at a position corresponding to
; OTHER INFORMATION: nucleotide 3107 of SEQ ID NO:1 as provided herein.
US-10-308-264-181

Query Match      23.8%; Score 398.8; DB 16; Length 16569;
Best Local Similarity 58.2%; Pred. No. 5e-110;
Matches 744; Conservative 0; Mismatches 522; Indels 12; Gaps 2;

QY 244 AACGGACCTGTGGAACGTCATGTTACTACCATGTTATCTCATGATGTTCTTTGTG 303
DB 5051 AACGACCAACATCTACACAGTTATGTCACAGCCCAAGCAATTTGTAATCTCTTCA 6110
QY 304 GGTATCCCGCGCATGTTGGTGTGTTGGTAACTATCTGATCCGCTCGAATCGCGCT 363
DB 6111 GTAATACCCATCATATATCGAGGCTTGGCAACTGACTAGTTCCCTAATATCGGTGCC 6170
QY 364 CCGGATATGCGCTTCCCGGTATGACACCTGTGTTCTGGCTGTTCTATTCGCGGTACC 423
DB 6171 CCGGATATGCGGTTTCCCGGCATAACAACAACTGCTTGAATCTTACCTCCCTCTCTC 6230
QY 424 GCGATGGGCGTGGCTTCGCTGTGTCGACCGGCGGTGACGGTCAGCTGGGTTCGGGCGTT 483
DB 6231 CTACTCTCTCGCATCTCTATATGAGGAGCGGAGCAGGAGGAGGTTGAACATGTC--- 6287
QY 484 GGTGGGTGTTGTACCCCGCGCTGTGCGACCCGGAGCTGGCTATTCGATGACCTCGG 543
DB 6288 -----TACCCTCCCTTAGCAGGAACTACTCCACCTCTGGAGCGCTCGTAGACCTAAC 6341
QY 544 ATTTTCGGGTTTCACTTGTGGGTGCTCTCGATCATGGGCGGATCAACATGATCAG 603
DB 6342 ATCTTCTCTTACACCTAGAGGTGTTCTCTTAATCTTAGGGCCATCAATTCATCA 6401
QY 604 ACCTTCTTGAACATGCGCGCCCGCGCATGACGCTGCAAAAGTGCGGTGTTCTCGTGG 663
DB 6402 ACAATATCAATATAAAACCCCTGCGCATAACCCCAATACCAACGCCCTCTTCTCTGA 6461
QY 664 TCGATCTTTATACGCTTGGCTGATCCTGCGGCGCTCGGCTCGGCTGCTGCTGCTGCAATC 723
DB 6462 TCGCTCTTAATACAGAGTCTTACTTCTCTTATCTCTCCAGTCTAGCTGCTGCTG 6521
QY 724 ACCATCTGTGTGACCGACGCTAACTTCGGCAGACCTTCTTCAATCTGCTGCGCGCGT 783
DB 6522 ACCATACTACTACAGACCGCAACCTCAACACACCTTCTTCGACCCCGCGGAGGAGGA 6581
QY 784 GACCCGATCTGTACCAACATCTCTGTTGTTTGGSCACCCGGAAGTGTACATCATC 843
DB 6582 GACCCCATCTTATACCAACACCTATTCTGATTTTTCGGTCAACCTGAAATTTATATCT 6641
QY 844 ATTCTGCGCGGTTGGCATCATCAGCCATGTCGTTGCGACCTTCTC---GAAAGCGCG 900
DB 6642 ATCTTACGAGCTTCGGAATATCTCCCATATTGTAACITACTACTCCGGAAGGAGAA 6701
QY 901 GTCTTCGTTACTGCGGATGTTCTATGCAATGGTGGCAATCGGTGTTCTGGCTTTGTC 960
DB 6702 CCATTGGATACATAGTATGGTCTGAGCTATGATATCAATTTGGCTTCTAGGTTTATC 6761

US-10-308-264-440
Query Match      23.8%; Score 398.8; DB 16; Length 16569;
Best Local Similarity 58.2%; Pred. No. 5e-110;
Matches 744; Conservative 0; Mismatches 522; Indels 12; Gaps 2;

QY 961 GTCTGGGCGACCAACATGATACCGTTGGTATGTCCTGCTGACCCAGCAATCTTACTTCATG 1020
DB 6762 GTGTGAGCACACCATATATTTACAGTAGGAATAGAGCTAGACACAGAGCATATTTTACC 6821
QY 1021 CTGGCCACCATGTTGATCGGGTGGCGACCGGCAATTAAGATCTTCTCGTGGATCGCCAGC 1080
DB 6822 TCGGCTACCAATATCATCGCTATCCCGCGGCTCAAGTATTTAGCTGACTCGCCACA 6881
QY 1081 ATGTTGGGCGGCTCGGTTGAGTTCAATATCCCGATGCTCTGGGCTTTGGCTTTATGTTTC 1140
DB 6882 CTCACGGAAGCAATATGAAATGATCTGCTGCAAGTCTCTGAGCCCTAGGATTCATCTTT 6941
QY 1141 CTGTTACCGTGGTGTGTGACCGGTATCGTCTGCGCCCAAGCGGTCTTGGACCCGTGCA 1200
DB 6942 CTTTTACCGTAGGTGGCTGACTGGCAATGTTATAGCAACTCATCACTAGACATCGTA 7001
QY 1201 TATCAGACACCTATTACGTGGTGGGCGCATCTCCATATATGATGTCGTGGGTGGATC 1260
DB 7002 CTACACGACAGTACTACGTTGTAGTCTCACTTCCACTATGCTCTATCAATAGGAGCTGA 7061
QY 1261 TTTGCGATCTTCGCGGTATCTACTTTTACATCGGAAGTTCTCGGCGCGGCTTTTCCCG 1320
DB 7062 TTTGCCATCATAGAGGCTTCATCTCACTGATTTCCCTATTTCTCAGGCTACACCTAGAC 7121
QY 1321 GAATGGGCTGCAAGCTGCACTTTCTGGACCTTCTTCATCGTGGAGCAAGTCACGTTCTTC 1380
DB 7122 CAACCTACGCGCAAAATCCATTTCACTATCATATTCATCGCGGTAAATCTAACTTTCTTC 7181
QY 1381 CCGGACGACCTTCCTGGGACGTCAGGATGTCGGCGGCTTACATCGACTATCCGAGACC 1440
DB 7182 CCACACACCTTTCTCGGCTTATCCGGAATGCGCGGAGTTACTCGGACTACCCCGATGCA 7241
QY 1441 TTCCGCTGTGGAAACAAAGTCTGCTCTATGTTGCTGCTTCTCGGCTTCTGCTCTG 1500
DB 7242 TACACACATGAACATCTCTATCATCTGTTAGGCTCAATCTTCTCTACAGAGTAATA 7301
QY 1501 TTCTTCATCGTATCTTT 1518
DB 7302 TTAATAATTTTTCATGATT 7319

RESULT 8
US-10-308-264-440
; Sequence 440, Application US/10308264
; Publication No. US20040029133A1
; GENERAL INFORMATION:
; APPLICANT: Herxinstadt, Corinna
; TITLE OF INVENTION: MITOCHONDRIAL DNA POLYMORPHISM
; FILE REFERENCE: 660088.451
; CURRENT APPLICATION NUMBER: US/10/308,264
; CURRENT FILING DATE: 2002-11-25
; NUMBER OF SEQ ID NOS: 697
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 440
; LENGTH: 16569
; <211> conflict
; TYPE: DNA
; ORGANISM: Homo sapiens mitochondrial
; FEATURE:
; LOCATION: 3107
; OTHER INFORMATION: "n" refers to a nucleotide at position 3107 of the
; OTHER INFORMATION: Cambridge Reference Sequence (CRS, Anderson et al.
; OTHER INFORMATION: 1981 Nature 290:457-465), later determined to be
; OTHER INFORMATION: absent from the human mtDNA sequence according to
; OTHER INFORMATION: the revised CRS (rCRS, Andrews et al. 1999 Nature
; OTHER INFORMATION: Genetics 23:147); "n" therefore represents a
; OTHER INFORMATION: single nucleotide deletion at a position corresponding to
; OTHER INFORMATION: nucleotide 3107 of SEQ ID NO:1 as provided herein.
US-10-308-264-440
```

244 AACGACACCTGTGAAAGCTCATGTTACCTTACCATGTTATCTTGATGATGTTCTTTTG 303
Db
6051 AACGACACATCTCAAGCTTATCGTCACAGCCCATGTTGTAATATCTTTTCATA 6110
Qy
304 GGTATCCCGCATGTTGCGGTGTTTGGTAACTATCTGATCGCGTGCAGAAATCGCGCT 363
Db
6111 GTAATACCCATCAATAATCGGAGGCTTTGGCAACTGACTAGTTCCTCCCTAAATTCGGTGCC 6170
Qy
364 CCGATATGGCTTTCCCGCGTATGAACAAACCTGTCGTTCTGGCTGTTTCATTTGCGCGGTACC 423
Db
6171 CCCGATATGGCTTTCCCGCATAAACACATAGCTTTGACTCTTACCTCCCTCTCTC 6230
Qy
424 CGATGGCGGTGGCTTCTGCTGTCGACCGGGCGGTGACGGTACGCTGGGTTCGGGGTT 483
Db
6231 CTACTCTGCTCGCATCTGCTATAGTGGAGCCCGGACGAAACAGGTTGAACAGTC--- 6287
Qy
484 GTTTGGGTTCTGTACCGCGCTCTCGACCGCGAAGCTGCTATTTCGATGGACCTCGCG 543
Db
6288 -----TACCTCTCTTAGCAGGAACTACTCCACCCCTGGAGCCTCGTAGACCTAAC 6341
Qy
544 ATTTTCGGGTTCATCTGTCGGGTGCTCTCTCGATCATGGCGCGATCAACATGATCAAG 603
Db
6342 ATCTTCTCTTACACCTAGCAGGTGTCTCTCTATCTTAGGGGCAATCAATTTTCATCACA 6401
Qy
604 ACCTTCTTGAACATGCGCGCCCGCATGACGCTGCACAAAGTGCCTGTTCTCGTG 663
Db
6402 ACAATTAATCAATAAACCCTTGCCTAACCCATACCAAGCCCTCTTCTGCTGA 6461
Qy
664 TCGATCTTTATCACGGCTTGGCTGATCTGCTGCGCTGCGGTTCGCGTGGTGAATC 723
Db
6462 TCCGTCCTAATCACAGCAGTCTACTTCTCTCTCTCTCCAGCTCTAGCTGCGCATC 6521
Qy
724 ACCATGCTGTGACGACCGTAATCTCGGCACGACCTTCTCAATCTGCTGCGCGGT 783
Db
6522 ACTATACTACTAAGACCGCAACTCAACACCACTTCTTCGACCCCGCGGAGGAGGA 6581
Qy
784 GACCCGATTCTGTACCAACACATCTCTGCTGTTCTTGGGACCGGAGGTATCATCATC 843
Db
6582 GACCCCATCTATACCAACACCTATCTGATTTTCGTCACCTGAGTTTATATCTT 6641
Qy
844 ATTCGCGCGCTTTGGCATCAACAGCCTGCTGCTGACCTTCTC-----AGCCG 900
Db
6642 ATCTTACCGCTTCGAATAATCTCCATATTGTAACTTACTACTCCGGAAGAAAGAA 6701
Qy
901 GTCTTCGTTACCTGCGATGGTCTATGCAATGTTGGCAATGCTGTTCTGCGCTTTGTC 960
Db
6702 CCATTTGGATACATAGTATGGTCTGACCTATGATATCAATTTGGCTTCTAGGGTTATC 6761
Qy
961 GTCTGGGCGACCATGTATACCGTTGGTATGTCGCTGACCGCAATCCTACTTCAATG 1020
Db
6762 GTGTGAGCACCATATATTTACAGTAGGAATAGACGTAGACACAGGACATATTTCAAC 6821
Qy
1021 CTGCGCCACCATGGTATCGGGTCCGACCGGATTAAGATCTTCTCGGATCGCCAG 1080
Db
6822 TCCGCTACCATATATCATCGCTATCCCGACCGGGTCAAGATTTAGCTGACTCGCCACA 6881
Qy
1081 ATGTGGGCGGCTCGGTGAGTTCAATCGCCGATGCTCTGGGCTTTGCTTTATGTTTC 1140
Db
6882 CTCACGGAAGCAATATGAATGATCTGCTGACGTGCTCTGAGCCCTAGGATTTATCTTT 6941
Qy
1141 CTGTTACCGTGGGTGATGACCGGTATGTCGCTGGCCCAAGCGGTCTGACCGGTGA 1200
Db
6942 CTTTTCACCGTAGGTGCGCTGACGTGGATTTGATTAGCAAACTCATCATAGACATCGTA 7001
Qy
1201 TATCACGACACCTATTACGTGGTGGCGCACTTCCATTTATGATGTCGTTGGGTGGATC 1260
Db
7002 CTACACGACACGCTACTAGTTGATGCTCACTTCCATATGCTCTATCAATAGAGCTGTA 7061
Qy
1261 TTTGCGATCTTCCCGGTATCTACTTTTACATGCGGAAGTTCTCGGGCGCGCTTTCCCG 1320
Db
7062 TTTGCCATCATAGGAGGCTTCATTTCACTGATTTTCCCTATTTCTCAGGCTACACCTTAGAC 7121

1321 GAATGGCTGCAAGCTGCACCTTCTGACCTTCTTCAATCGTGCGAAGCTCACGTTCTTC 1380
Db
7122 CAAGCTACGCGCAAAATCCATTTCACTATCATATCAATCGCGGTAATCTAACTTTCTTC 7181
Qy
1381 CCGAGCACTTCTCTGGGAGCTCAGGTTATCGCGGCTTACATCGACTATCCCGAGGCC 1440
Db
7182 CCACAACTTCTCTCGGCTATCCGGAATGCCGACGTTACTCGGACTACCCGATGCA 7241
Qy
1441 TTCGCGCTGTGGAACAAAGTCTCGTCTATGTTGCTTCTGGCTTCGCGCTTCGCTCGT 1500
Db
7242 TACACCATGAACATCTATCATCTGTAGGCTCATTTCTTAACAGCAGTAATA 7301
Qy
1501 TTCCTCATCGTATCTTT 1518
Db
7302 TTAATAATTTTCATGATT 7319
RESULT 9
US-10-308-264-469
Sequence 469, Application US/10308264
Publication No. US20040029133A1
GENERAL INFORMATION:
APPLICANT: HerinStadt, Corinna
TITLE OF INVENTION: MITOCHONDRIAL DNA POLYMORPHISM
FILE REFERENCE: 660088.461
CURRENT APPLICATION NUMBER: US/10/308.264
CURRENT FILING DATE: 2002-11-25
NUMBER OF SEQ ID NOS: 697
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 469
LENGTH: 16569
TYPE: DNA
ORGANISM: Homo sapiens mitochondrial
FEATURE:
LOCATION: 3107
OTHER INFORMATION: "n" refers to a nucleotide at position 3107 of the
OTHER INFORMATION: Cambridge Reference Sequence (CRS, Anderson et al.
OTHER INFORMATION: 1981 Nature 290:457-465), later determined to be
OTHER INFORMATION: absent from the human mtDNA sequence according to
OTHER INFORMATION: the revised CRS (rCRS, Andrews et al. 1999 Nature
OTHER INFORMATION: Genetics 23:147); "n" therefore represents a
OTHER INFORMATION: single nucleotide deletion at a position corresponding to
OTHER INFORMATION: nucleotide 3107 of SEQ ID NO:1 as provided herein.
US-10-308-264-469

Query Match 23.8%; Score 398.8; DB 16; Length 16569;
Best Local Similarity 58.2%; Pred. No. 5e-110;
Matches 744; Conservative 0; Mismatches 522; Indels 12; Gaps 2;
Qy 244 AACGACACCTGTGGAAGCTCATGTTACCTTACCATGTTATCTTGATGATGTTCTTTTG 303
Db 6051 AACGACACATCTCAAGCTTATCGTCACAGCCCATGCTTTGTAATATCTTTTCATA 6110
Qy 304 GGTATCCCGCATGTTGCGGTGTTTGGTAACTATCTGATCGCGCTGCAATCGCGCT 363
Db 6111 GTAATACCCATCAATAATCGGAGGCTTTGGCAACTGACTAGTTCCCTTAATTCGGTGCC 6170
Qy 364 CCGATATGGCTTTCCCGCGTATGAACAAACCTGCTGCTTCTGGCTGTTTCATTTGCGCGTACC 423
Db 6171 CCCGATATGGCTTTCCCGCATAAACACATAGCTTCTGACTCTTACCTCCCTCTCTC 6230
Qy 424 CGATGGCGGTGGCTTCTGCTGTCGACCGGGCGGTGACGGTACGCTGGGTTCGGGGTT 483
Db 6231 CTACTCTGCTCGCATCTGCTATAGTGGAGCCCGGACGAAACAGGTTGAACAGTC--- 6287
Qy 484 GTTTGGGTTCTGTACCGCGCTCTCGACCGCGAAGCTGCTATTTCGATGGACCTCGCG 543
Db 6288 -----TACCTCTCTTAGCAGGAACTACTCCACCCCTGGAGCCTCGTAGACCTAAC 6341
Qy 544 ATTTTCGGGTTCATCTGTCGGGTGCTCTCTCGATCATGGCGGATCAACATGATCAAG 603
Db 6342 ATCTTCTCTTACACCTAGCAGGTGTCTCTCTATCTTAGGGCCATCAATTTTCATCACA 6401

604 ACCTCTTGAACATGGCGCGCCCGGCGATGACGCTGCACAAAGTGGCGTGTCTTCGTGG 663
Db ACAAATATCAATATAAAACCCCGCTGCATAAACCAATACCAACGCCCTCTTCGTGTA 6461
QY TCGATCTTTATCAGCGCTTGGTGTCTTCCTGCTGGGCTGCGGCTCTGGCTGGTCAATC 723
Db TCGGCTCTAATCACAGCAGTCTACTTCTCTCTATCTCTCCAGTCTAGTCTGGCATC 6521
QY ACATGCTCTGACCGACCGTAACTTCGCGCAGACCTTCTCAATCTGCTGGCGCGGT 783
Db ACTATACTACTAACAGACCGCAACCTCAACACCACTCTTCGACCCCGCGGAGGGA 6581
QY GACCGCATCTGTACCAACACATCTGTGTCTTGGCACCGGAGTGTACATCATC 843
Db GACCCCATCTTATACCAACACCTTCTGTATTTTGGTACCCCTGAAGTTTATATCTT 6641
QY ATTCTGCCCGCTTGGCATCATCAGCCATGCTGTGCACTTCTC---GAAAGGCG 900
Db ATCTTACCAGGCTTCGGAATAATCTCCCATATTGAATTTACTTCTCCGGAAGAA 6701
QY GTCTTCGGTACTGCGGATGCTATGCAATGGTGGCAATCGTGTCTGGGCTTCTC 960
Db CCATTTGGATACATAGTATGCTGTAGCTATGATATCAATTTGGTCTCTAGGTTTATC 6761
QY GTCTGGCGCACACATGTACACCGTGTGTATGTCTGCTGACCCAGCAATCTTCTCATG 1020
Db GTGTAGCACACCATATATTACAGTAGGAATAGAGTAGACACAGCATATTTCACC 6821
QY CTGGCCACCATGTTGATCGGGTGGCGGACCGGATTAAGATCTTCTGTGGATGCGCAG 1080
Db TCGCTTACCATAATCATCTGCTATCTCCACCGCGTCAAAAGTATTTAGCTGACTGCCACA 6881
QY ATGTGGGGCGGCTCGGTTCAGTTTCAATCGCGATGCTCTGGGCTTTTGGCTTTATGTC 1140
Db CTCCACGGAGCAATATGAATGATCTGTGAGTGTCTGTAGCCCTAGGATTCATCTT 6941
QY CTGTTACCGTGGTGTGTGACCGGTATGCTGTGCTGGCCCAAGCGGCTGGAACGTTGA 1200
Db CTCTTACCGTAGTGGCTGACTGGCATTTATTAGCAAACTCATCATAGACATCGTA 7001
QY TATCAGACACCTATTACGTTGGTGGCGCACTTCCATTTATGATGTGCTGGTGGTGGATC 1260
Db CTACAGACACGTAATGCTGTGGTGGCGCACTTCCACTTCCACTTGTCTTCAATAGAGTGA 7061
QY TTTCGATCTTCCCGGTATCTACTTTTACATGCGCAAGTTCTCGGGCGCGCTTTCCCG 1320
Db TTTCGATCATAGAGGCTTCTACTTCTGATTTCCCTTATCTCAGGCTACACCTTAGAC 7121
QY GAATGGCTGCAAGTGTGATCTGACCTTCTGACCTTCTGATGGTGGCACTGACGTTCTTC 1380
Db CAAACCTACGCCCAAAATCCATTTCACTATCATATTATCAGCGTAAATCTAACTTTCTTC 7181
QY CCGCAGCACTTCTCTGGGAGCTCAGGCTATGCGCGCGCTTACATCGACTATCCCGAAGCC 1440
Db CGACACACTTCTCGGCTATCCGGAATGCCGAGCTTACTCGGACTACCCGATGCA 7241
QY TTCGCGTGTGGAACAAAGTCTGCTCTCTATGTTGGTGGCTTCTGCGCTTCCGCTGCTTC 1500
Db TACACACATGAAACATCTCTATCATCTGTAGGCTCATCTATTCTCTAAACAGCAGTAATA 7301
QY TTCTTCATCTGATCTTT 1518
Db TTAATAATTTTCATGATT 7319

RESULT 10

US-10-308-264-474
; Sequence 474, Application US/10308264
; Publication No. US20040029133A1
; GENERAL INFORMATION:
; APPLICANT: Herinstad, Corinna
; TITLE OF INVENTION: MITOCHONDRIAL DNA POLYMORPHISM

FILE REFERENCE: 660088.461
CURRENT APPLICATION NUMBER: US/10/308,264
CURRENT FILING DATE: 2002-11-25
NUMBER OF SEQ ID NOS: 697
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 474
LENGTH: 16569
<211> conflict
TYPE: DNA
ORGANISM: Homo sapiens mitochondrial
FEATURE:
LOCATION: 3107
OTHER INFORMATION: "n" refers to a nucleotide at position 3107 of the
OTHER INFORMATION: Cambridge Reference Sequence (CRS, Anderson et al.
OTHER INFORMATION: 1981 Nature 290:457-465), later determined to be
OTHER INFORMATION: absent from the human mtDNA sequence according to
OTHER INFORMATION: the revised CRS (rCRS, Andrews et al. 1999 Nature
OTHER INFORMATION: Genetics 23:147); "n" therefore represents a
OTHER INFORMATION: single nucleotide deletion at a position corresponding to
OTHER INFORMATION: nucleotide 3107 of SEQ ID NO:1 as provided herein.
US-10-308-264-474

Query Match 23.8%; Score 398.8; DB 16; Length 16569;
Best Local Similarity 58.2%; Pred. No. Se-110;
Matches 744; Conservative 0; Mismatches 522; Indels 12; Gaps 2;

QY 244 AACGGACACCTGTGGAAACGTCATGTTACCTTACCATTGATGATGATGATGATGATG 303
Db 6051 AACGACACATCTACAACGTTATCGTCACAGCCCATGCAITTTGTAATAATCTTTCATA 6110
QY 304 GGTATCCCGGCAATTTTCGGTGGTTTGGTAACTATCTGATGCGCTGCAATCGGCGCT 363
Db 6111 GTAATACCCATCAATAATCGAGGCTTTGGCAACTGACTAGTTCCTCCCTAATAATCGGTGCC 6170
QY 364 CCGGATATGGCTTCCCGCGTATGAACAACCTGTCTGTCTGTGGCTGTTCATTCGCCGTACC 423
Db 6171 CCGATATGGGTTTCCCGCATTAACAATAGCTTCTGACTTTACCTCCCTCTCTC 6230
QY 424 CGGATGGCGTGGTTCGCTGTTCCGACCGGCGGTGACGCTGAGTGGTGGGCGGT 483
Db 6231 CTACTCTGCTCGCATCTGCTATAGTGGAGCGCGAGCAGGTAACAGTCTGAACAGTC--- 6287
QY 484 GGTGGGTTCTGATACCGCGCTCTGACCGCGCAAGCTGCTATTCGATGGACCTCGCG 543
Db 6288 -----TACCTCTCCTTAGCAGGGAATCTATCCACCCCTGAGCTCTCGTAGACTAACC 6341
QY 544 ATTTTCGGGTTCACTTGTTCGGGTGCTCTCTCGATCATGCGCGCGATCAACATGATCAG 603
Db 6342 ATCTCTCTTACACCTAGCAGGTTCTCTCTCTATCTTAGGGCCATCAATTTTCATCACA 6401
QY 604 ACCTTCTTGAACATGCGGCGCGCGGATGACGCTGCACAAAGTGGCGTGTCTTCGTGG 663
Db 6402 ACAATTATCAATAAACCCTTGCATTAACCAATACCAACCGCCCTCTTCGTCTGA 6461
QY 664 TCGATCTTATCAGCGTTGGCTGATCTGCTGGCGCTGCGGTTCTGGCTGGTGGCAATC 723
Db 6462 TCGTCTTATACAGCAGTCTTACTTCTCTATCTCTCCAGTCTAGCTGGCATC 6521
QY 724 ACCATGCTGTGACGCGCGGTAATTCGGACAGACCTTCTTCAATCTGCTGGCGCGGT 783
Db 6522 ACTATACTACTAACAGACCGCAACCTCAACACCACTTCTTCGACCCCGCGGAGGGA 6581
QY 784 GACCGGATCTGTACCAACACATCTGTGGTCTTTTGGGACCCCGGAAGTGTACATCATC 843
Db 6582 GACCCCATCTTATACCAACACCTATTCTGATTTTTCGCTACCTGAGTTTATATCTT 6641
QY 844 ATCTGCCCGCTTGGCATCATCAGCAATGCTGTGCGACTTCTC---GAAAGGCG 900
Db 6642 ATCTTACCAGGCTTCGGAATAATCTCCCATATTGTAATTTACTACTTCCGGAAGAA 6701
QY 901 GTCTTCGCTTACCTGCCGATGCTATGCAATGTTGGCAATCGGTTCTCGGCTTGTG 960
Db 6702 CCATTTGGATACATAGTATGCTGTAGCTATGATATCAATTTGGTCTCTAGGTTTATC 6761
QY 961 GTCTGGCGCACACATGTACACCGTGTGTATGTCTGCTGACCCAGCAATCTTCTCATG 1020
Db 6762 GTGTAGCACACCATATATTACAGTAGGAATAGAGTAGACACAGCATATTTCACC 6821
QY 1021 CTGGCCACCATGTTGATCGGGTGGCGGACCGGATTAAGATCTTCTGTGGATGCGCAG 1080
Db 6822 TCGCTTACCATAATCATCTGCTATCTCCACCGCGTCAAAAGTATTTAGCTGACTGCCACA 6881
QY 1081 ATGTGGGGCGGCTCGGTTCAGTTTCAATCGCGATGCTCTGGGCTTTTGGCTTTATGTC 1140
Db 6882 CTCCACGGAGCAATATGAATGATCTGTGAGTGTCTGTAGCCCTAGGATTCATCTT 6941
QY 1141 CTGTTACCGTGGTGTGTGACCGGTATGCTGTGCTGGCCCAAGCGGCTGGAACGTTGA 1200
Db 6942 CTCTTACCGTAGTGGCTGACTGGCATTTATTAGCAAACTCATCATAGACATCGTA 7001
QY 1201 TATCAGACACCTATTACGTTGGTGGCGCACTTCCATTTATGATGTGCTGGTGGTGGATC 1260
Db 7002 CTACAGACACGTAATGCTGTGGTGGCGCACTTCCACTTCCACTTGTCTTCAATAGAGTGA 7061
QY 1261 TTTCGATCTTCCCGGTATCTACTTTTACATGCGCAAGTTCTCGGGCGCGCTTTCCCG 1320
Db 7062 TTTCGATCATAGAGGCTTCTACTTCTGATTTCCCTTATCTCAGGCTACACCTTAGAC 7121
QY 1321 GAATGGCTGCAAGTGTGATCTGACCTTCTGACCTTCTGATGGTGGCACTGACGTTCTTC 1380
Db 7122 CAAACCTACGCCCAAAATCCATTTCACTATCATATTATCAGCGTAAATCTAACTTTCTTC 7181
QY 1381 CCGCAGCACTTCTCTGGGAGCTCAGGCTATGCGCGCGCTTACATCGACTATCCCGAAGCC 1440
Db 7182 CGACACACTTCTCGGCTATCCGGAATGCCGAGCTTACTCGGACTACCCGATGCA 7241
QY 1441 TTCGCGTGTGGAACAAAGTCTGCTCTCTATGTTGGTGGCTTCTGCGCTTCCGCTGCTTC 1500
Db 7242 TACACACATGAAACATCTCTATCATCTGTAGGCTCATCTATTCTCTAAACAGCAGTAATA 7301
QY 1501 TTCTTCATCTGATCTTT 1518
Db 7302 TTAATAATTTTCATGATT 7319

QY 961 GTCTGGGCGACCAATGATACCGTTGGTATGCTGAGCCAGCAATCTTACTTCATG 1020
 Db 6762 GTGTGAGCACACCAATATTTACAGTAGGAATAGACGACGACATATTTTCAACC 6821
 QY 1021 CTGGCCACCATGGTATGCGGGTCCGACCGGCAATTAAGATCTTCTCGTGATCGGCAG 1080
 Db 6822 TCCGCTACCAATAATCATCGGTATCCCGACCGGGTCAAGATATTTAGCTAGCTCGGCACA 6881
 QY 1081 ATGTGGGCGCGCTCGGTGAGTTCAAAATCGCGATGCTCTGGGCGTTTGGCTTTATGTTT 1140
 Db 6882 CTCACGGAAGCAATATGAATGATCTGCGAGTCTCTGAGCCCTAGGATCACTTT 6941
 QY 1141 CTGTTACCGTGGTGTGACCGGTATCGTGTGCGCCCAAGCGGTCTGGAACCTGCA 1200
 Db 6942 CTTTTCACCGTAGTGGCTGACTCGGCAATTTGATAGCAAACTCATCACTAGACATCGTA 7001
 QY 1201 TATCACGACACCTATTACGTTGGCGCACTTCCATTATGATGCTCGCTGGGTGGATC 1260
 Db 7002 CTACGACACGTAAGTATGTTGAGCCACTTCCACTATGCTCTATCAATAGAGCTGTA 7061
 QY 1261 TTTGCGATCTTCCGCGTATCTATTTTATCATGCCGAAGTTCTCGGCGCGCTTTCCCG 1320
 Db 7062 TTTGCCATCATAGGAGGCTTCATTCACTGATTTCCCTATTTCTCAGGCTACACCCCTAGAC 7121
 QY 1321 GAATGGGCTGCAAGCTGCACTTCTGACCTTCTTCATCGGTGCGAAGCTCACTGTTCTT 1380
 Db 7122 CAACCTAGCCAAATCCATTTCACATCATATCATGCGGTAAATCTTAACTTTCTTTC 7181
 QY 1381 CCGCAGCACTTCTGGGACGTCAGGATGCGCGCGCTTACATCGACTATCCGGAAGC 1440
 Db 7182 CCAACACATTTCTCGGCTATCCGGAATGCGCGAGCTTACTCGGACTACCCCGATGCA 7241
 QY 1441 TTGCGGTGGAAGTCTGCTATGTTGCGGTCTGCGCTTCGCTCGCTCGTTCCTG 1500
 Db 7242 TACACCAATGAACATCCCTATCATCTGATGCTAGGCTCATTTCTTCAACAGCAGTAATA 7301
 QY 1501 TTTCTCATCTGATCTTT 1518
 Db 7302 TTAATAATTTTCATGATT 7319
 RESULT 11
 US-10-308-264-476
 ; Sequence 476, Application US/10308264
 ; Publication No. US20040029133A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hermsstadt, Corinna
 ; TITLE OF INVENTION: MITOCHONDRIAL DNA POLYMORPHISM
 ; FILE REFERENCE: 660088.461
 ; CURRENT APPLICATION NUMBER: US/10/308,264
 ; CURRENT FILING DATE: 2002-11-25
 ; NUMBER OF SEQ ID NOS: 697
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 476
 ; LENGTH: 16569
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens mitochondrial
 ; FEATURE:
 ; LOCATION: 3107
 ; OTHER INFORMATION: "n" refers to a nucleotide at position 3107 of the
 ; OTHER INFORMATION: Cambridge Reference Sequence (CRS, Anderson et al.
 ; OTHER INFORMATION: 1981 Nature 290:457-465), later determined to be
 ; OTHER INFORMATION: absent from the human mtDNA sequence according to
 ; OTHER INFORMATION: the revised CRS (rCRS, Andrews et al. 1999 Nature
 ; OTHER INFORMATION: Genetics 23:147); "n" therefore represents a
 ; OTHER INFORMATION: single nucleotide deletion at a position corresponding to
 ; OTHER INFORMATION: nucleotide 3107 of SEQ ID NO:1 as provided herein.
 US-10-308-264-476

Query Match 23.8%; Score 398.8; DB 16; Length 16569;
 Best Local Similarity 58.2%; Pred. No. 5e-110;

Matches 744; Conservative 0; Mismatches 522; Indels 12; Gaps 2;
 QY 244 AACGACACCTGTGGACGTCATGTTTACCTTACCATGTTATTTCTGATGATGTTTCTTTGTG 303
 Db 6051 AACGACACATCTACACGTTTATCGTCACAGCCCATGATTTGTAATAATCTTCTTCATA 6110
 QY 304 GGTATCCCGCATGTTTCTGGTGGTTTGGTAACTATCTGATGCGCGTGCATAATCGGCGCT 363
 Db 6111 GTAATACCCATCATATAATCGAGGCTTTGGCAACTGACTAGTTCCCTTAATAATCGGTGCC 6170
 QY 364 CGGATATGGCTTCCCGCTATGAACAACTGCTGCTTCTGGCTGTTTCATTGCCCGTACC 423
 Db 6171 CCGGATATGGGTTTCCCGCATTAACAACTAAGCTTCTGACTTACCTTCCCTCTCTC 6230
 QY 424 GCGATGGCGGTGGTCTGCTGTCGACCGGGGGTACGCTCAGCTGGGTTCGGGGGTT 483
 Db 6231 CTACTCTCTGCTCGCATCTGCTATAGTGAGGCGCGAGCAGGAACAGGTTGAACAGTC --- 6287
 QY 484 GGTGGGTGCTGTACCCCGCTGTCGACCGCGCTGCTGCTGCTATTCGATGCACTCGCG 543
 Db 6288 -----TACCTCTCCCTTAGCAGGGAACCTACTCCACCTCGAGGCTCGTAGACTAAC 6341
 QY 544 ATTTTCGCGTTACTTGTGCGGTGCTTCTGATCATGTCGCGCGATCAACATGATCAAG 603
 Db 6342 ATCTTCTCTTACACCTAGCAGGTGCTCTCTATCTTAGGGGCGCATCAATTTCAATCACA 6401
 QY 604 ACCTTCTTGAACATGCGCGCGCGCGGATGAGCTGACAAAGTGCCTGTTGTTCTCGTGG 663
 Db 6402 ACAATATCAATATAAAACCCCTGCCATAACCAATACCAACGCGCTCTTCTGTCGA 6461
 QY 664 TCGATCTTTATACGCGTTTGGCTGATCCTGCTGCGCTGCGCGTTCTGGTGGTGCATC 723
 Db 6462 TCGCTCTTAATACAGCAGTCTCTACTTCTCTCTCTCTCTCCAGCTCTAGTCTGCGCATC 6521
 QY 724 ACCATGCTGTCGCGACCGTAACTTCGGCAGACCTTCTTCAATCTGCTGGCGGGGT 783
 Db 6522 ACTATACTAAACAGACCGCAACCTCAACACCACTTCTTCGACCCCGCGGAGGAGGA 6581
 QY 784 GACCGGATTCGTACCAACACATCCTGTGGTTCCTTGGGACCCCGGAAGTGTACATCATC 843
 Db 6582 GACCCATCTATACCAACACTTATCTGATTTTTCGTCACCTCGAAGTTTATTTCTT 6641
 QY 844 ATTCGCGCGCTTTGGCATCATCAGCATGCTGTGTCGACCTTCTC ---GAAAAAGCG 900
 Db 6642 ATCTTACCCAGGCTTCGGAATTAATCTCCCATATTTAACTTACTACTCCGGAAGAAAGAA 6701
 QY 901 GTCTTCGGTTACCTGCGGATGCTATGCAATGTTGGCAATCGGTGTTCTGGGCTTTGTC 960
 Db 6702 CCAATTTGATACATAGGATGCTGAGCTATGATATCAATTTGGCTTCTTAGGTTTATC 6761
 QY 961 GTCTGGGCGCACCATGATACACCGTTGGTATGTCGTGACCCAGCAATCTTACTTCATG 1020
 Db 6762 GTGTGAGCACACCATATTTACAGTAGGAATAGACGCTAGACACAGCATATTTTCAAC 6821
 QY 1021 CTGGCCACCATGTTGATCGCGTCCGACCGGCAATTAAGATCTTCTGTCGATGCGCACG 1080
 Db 6822 TCCGCTACCAATAATCATCGGTATCCCGACCGGTCTCAAGTATTTAGCTAGCTCGGCACA 6881
 QY 1081 ATGTGGGCGCGCTCGGTGAGTTTCAAAATCGCGGATGCTCTGGGCGTTTGGCTTTATGTT 1140
 Db 6882 CTCACGGAAGCAATATGAATGATCTGCTGAGCTGCTGAGCCCTAGGATTCATCTTT 6941
 QY 1141 CTGTTACCGTGGTGTGACCGGTATCGTGTGCGCCCAAGCGGTCTGGAACCTGCA 1200
 Db 6942 CTTTTCACCGTAGTGGCTGACTCGGCAATTTGATAGCAAACTCATCACTAGACATCGTA 7001
 QY 1201 TATCACGACACCTATTACGTTGGCGCACTTCCATTATGATGCTCGCTGGGTGGATC 1260
 Db 7002 CTACGACACGTAAGTATGTTGAGCCACTTCCACTATGCTCTATCAATAGAGCTGTA 7061
 QY 1261 TTTGCGATCTTCCGCGTATCTATTTTATCATGCCGAAGTTCTCGGCGCGCTTTCCCG 1320
 Db 7062 TTTGCCATCATAGGAGGCTTCATTCACTGATTTCCCTATTTCTCAGGCTACACCCCTAGAC 7121

QY 1321 GAATGGGTCGCAAGTGTGACCTTCTGGACCTTCTTCAATCGGTGCGAAGTCAAGTTCCTTC 1380
DB 7122 CAAACCTACGCCCAAAATCCATTTCATCATATTCATCGCGGTAAATCTAACTTCCTTC 7181
QY 1381 CGCAGCAGCTTCTCGGAGCTGAGGTATGCGCGCGCTTACATCGATATCCCGAAGCC 1440
DB 7182 CCACAACACTTCTCGGCTATCCGGAATGCCCGACGTTACTCGGACTTACCCCGATGCA 7241
QY 1441 TTCGCGCTGTGAACAAAGTCTCGTCTATGTTGGGTTCTCGGCTTGGCTTCGTTCTGT 1500
DB 7242 TACACACATGAACATCTCATCTGTAGCTCATTCATTCTCTAACAGCAGTAATA 7301
QY 1501 TTCCTCATCGTGAATCTTT 1518
DB 7302 TTAATAATTTTCATGATT 7319

RESULT 12
US-10-308-264-501
; Sequence 501, Application US/10308264
; Publication No. US20040029133A1
; GENERAL INFORMATION:
; APPLICANT: Herrin, Corinna
; TITLE OF INVENTION: MITOCHONDRIAL DNA POLYMORPHISM
; FILE REFERENCE: 660088.461
; CURRENT APPLICATION NUMBER: US/10/308,264
; NUMBER OF SEQ ID NOS: 697
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 501
; LENGTH: 16569
; <211> conflict
; TYPE: DNA
; ORGANISM: Homo sapiens mitochondrial
; FEATURE:
; LOCATION: 3107
; OTHER INFORMATION: "n" refers to a nucleotide at position 3107 of the
; OTHER INFORMATION: Cambridge Reference Sequence (CRS, Anderson et al.
; OTHER INFORMATION: 1981 Nature 290:457-465), later determined to be
; OTHER INFORMATION: absent from the human mtDNA sequence according to
; OTHER INFORMATION: the revised CRS (rCRS, Andrews et al. 1999 Nature
; OTHER INFORMATION: Genetics 23:147); "n" therefore represents a
; OTHER INFORMATION: single nucleotide deletion at a position corresponding to
; OTHER INFORMATION: nucleotide 3107 of SEQ ID NO:1 as provided herein.
US-10-308-264-501

Query Match 23.8%; Score 398.8; DB 16; Length 16569;
Best Local Similarity 58.2%; Pred. No. 5e-110;
Matches 744; Conservative 0; Mismatches 522; Indels 12; Gaps 2;
QY 244 AACGGACACCTGTGGAACTGATGTTTACCTACATGGTATCTGATGATGTTCTTTGTG 303
DB 6051 AACGACCAATCTACAAAGTTATCGTCACAGCCATGCTTTGTAATAATCTTTCATA 5110
QY 304 GGTATCCCGCATGTTTCGGTGTGTTGTTGTAATCTATCTGATCGCGCTGAAATCGCGCT 363
DB 6111 GTAATACCATCATATAATCGGAGCTTTGGCACTGACTAGTTCCCTTAATAATCGGTGCC 6170
QY 364 CCGGATATGCTTCCGGGTATGAACACTCTGCTTCTGGGTGTTCAATCGCGTACC 423
DB 6171 CCGGATATGCGGTTTCCCGCATAAACATGAAGCTTCTGACTTCACTCCCTCTCTC 6230
QY 424 GCGATGGCGGTGCTTCTGCTGTTTCGACCGGGCGGTGACGGTCAAGTGTGGGCTTCGGGCGTT 483
DB 6231 CTACTCTGCTGCACTGCTATAGTGGAGGCGCGGAGCAAGGTTGAACAGTCT--- 6287
QY 484 GGTGGTGTCTGACCGCGCTGTGACCCCGGAAGCTGGCTATTTCGATGACCTCGCG 543
DB 6288 -----TACCCTCCCTTAGCAGGAACTACTCCACACCTGGAGCCCTCCGTAGACTAACCC 6341
QY 544 ATTTTCGCGGTTCACTGTGCGGTGCTCTCTGATCATGCGGCGCATCAACATGATCAGC 603

RESULT 13
US-10-308-264-508
; Sequence 508, Application US/10308264
; Publication No. US20040029133A1
; GENERAL INFORMATION:
; APPLICANT: Herrin, Corinna

DB 6342 ATCTTCTCCTTACACCTAGCAGGTGTCCTCTATCTTAGGGCCATCAATTTTCATCACA 6401
QY 604 ACCTTCTTGAACATGCGCGCCCGCGCATGACGCTGCACAAAGTGCCTTGTCTTCGTGG 663
DB 6402 ACAATATCAATAFAAAACCCCTTGCCTAATACCAATACCAACGCCCTCTTCGTCTGA 6461
QY 664 TCGATCTTTATCAGCGCTTGGCTGATCTCTGCTGGCGCTGCGGTTCTGGGCTGGTGAATC 723
DB 6462 TCCGCTCTAATACAGCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 6521
QY 724 ACCATGCTGTGACCGACCGTAACTTCGGCACACACCTTCTTCAATCTCTGTCGGCGCGGT 783
DB 6522 ACTATCTACTACAGACCGCACTCAACACACCTTCTTCGACCCCGCGGAGGAGGA 6581
QY 784 GACCGGATCTGTACCAACACATCTCTGTGTTCTTTGGGACCCCGGAAGTGTATCATCATC 843
DB 6582 GACCCCATTTCTATACCAACACCTATTCTGATTTTTCGCTCACCTGAGTTTATATCTT 6641
QY 844 ATTCTGCCGCTTGTGGCATCATCAGCCATGTCGTGTCGACCTTCTC---GAAAGACCG 900
DB 6642 ATCTTACAGGCTTCGGAATAATCTCCATATTGTAACTTACTCTCGGAAAAAAGAA 6701
QY 901 GTCTTGGTTTACCTGCCGATGGTCTATGCAATGGTGGCAATCGGTGTTCTGGGCTTTGTC 960
DB 6702 CCAATTTGATACATAGGTATGGTCTGAGCTATGATATCAATTTGGCTTCTTAGGCTTATC 6761
QY 961 GTCTGGGGGACCATGATGACCGTTGGTGTATGTCGTGACCCAGCAATCTCTACTTTCATG 1020
DB 6762 GTGTGAGCACACCATATATTTTACAGTAGGAATAGACGTAGACACACGAGCATATTTTACC 6821
QY 1021 CTGGCCACCATGATGTCGGGTGCCGACCGGCAATTAAGATCTTCTCGTGATCGCCACG 1080
DB 6822 TCCGCCACCAATATCATCGCTATCCCAACCGCGTCAAGTATTTAGCTGACTCGCCACA 6881
QY 1081 ATGTGGGGCGCTCGTTGAGTTCAAATCGCGATGCTCTGGGCTTGGCTTTATGTTTC 1140
DB 6882 CTCACGGAAGCAATATGAATGATCTGCTGAGTCTCTGAGCCCTAGGATTCATCTTT 6941
QY 1141 CTGTTTACCGTGGGTGTGTCACCGGTATCTGCTGGCCCAAGCGGTCTGGACCGTGCA 1200
DB 6942 CTTTTCACGTTAGTGGCTGACTGGCATGTTATAGCAAACTCATCACTAGACATGTA 7001
QY 1201 TATCAGACACCTATATGCTGGTGGCGACTTCCATATGTGATGTCGTGGGTGGGATC 1260
DB 7002 CTACACGACAGTACTACTAGTTGTAGCCCACTTCCACTATGTCTCTATCAATAGGAGCTGTA 7061
QY 1261 TTTGCGATCTTTCGCGGTATCTATCTTTTACATGCGGAAGTCTCGGCGCGCTTTCCTCG 1320
DB 7062 TTTGCCATCATAGAGGCTTCATTCATGATTTCCCTATTTCTCAGGCTACACCTAGAC 7121
QY 1321 GAATGGGTCGAAAGCTGCATCTTCTGGACCTTCTTCAATCGGTGCGAAGTCAAGTTCCTTC 1380
DB 7122 CAAACCTAGCGCAAAATCCATTTCACTATCATATTCATCGCGTAAATCTAACTTCTTCTC 7181
QY 1381 CCGGACACTTCTCGGACGTCAGGATGCGCGCGGTATACATCGACTATCCCGAAGCC 1440
DB 7182 CCACAACACTTCTTCGCGCTATCCGGAATGCCCGACGTTACTCGGACTACCCCGATGCA 7241
QY 1441 TTCGCGCTGTGAACAAAGTCTCGTCTCTATGTTGGGCTTCTCGGCTTCGCTTCCTG 1500
DB 7242 TACACCATGAACATCTCTATCATCTGTAGGCTCATTTCTCTAACAGCAGTAATA 7301
QY 1501 TTTTCATCGTGAATCTTT 1518
DB 7302 TTAATAATTTTCATGATT 7319

; TITLE OF INVENTION: MITOCHONDRIAL DNA POLYMORPHISM

; FILE REFERENCE: 660088.461

; CURRENT APPLICATION NUMBER: US/10/308,264

; CURRENT FILING DATE: 2002-11-25

; NUMBER OF SEQ ID NOS: 697

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 508

; LENGTH: 16569

; <211> conflict

; TYPE: DNA

; ORGANISM: Homo sapiens mitochondrial

; FEATURE:

; LOCATION: 3107

; OTHER INFORMATION: "n" refers to a nucleotide at position 3107 of the

; OTHER INFORMATION: Cambridge Reference Sequence (CRS, Anderson et al.

; OTHER INFORMATION: 1981 Nature 290:457-465), later determined to be

; OTHER INFORMATION: absent from the human mtDNA sequence according to

; OTHER INFORMATION: the revised CRS (rCRS, Andrews et al. 1999 Nature

; OTHER INFORMATION: Genetics 23:147); "n" therefore represents a

; OTHER INFORMATION: single nucleotide deletion at a position corresponding to

; OTHER INFORMATION: nucleotide 3107 of SEQ ID NO:1 as provided herein.

US-10-308-264-508

Query Match 23.8%; Score 398.8; DB 16; Length 16569;

Best Local Similarity 58.2%; Pred. No. 5e-110;

Matches 744; Conservative 0; Mismatches 522; Indels 12; Gaps 2;

244 AACGACACCTGTGGACGTCATGGTTACCTACCATGGTATCTGTGATGTTCTTTGG 303

6051 AACGACACATCTACACGTTATCGTCACAGCCCATTCATTTGTAATATCTTTTCA 6110

304 GGTATCCCGCATTTGTCGGTGGTTTGGTAACATATCTGATGCGGTGCAAAATCGGCGCT 363

6111 GTAATCCCATCAATATCGGAGGCTTTGGCAACTGACTAGTTCCTCAATATCGGTGCC 6170

364 CCGGATATGGCTTCCGGGTATGAACAACCTGTGTTCTGGTGTTCATATGCGGTAC 423

6171 CCGGATATGGCTTCCGGGTATGAACAACCTGTGTTCTGGTGTTCATATGCGGTAC 5230

424 CGGATGGGCGTGGTTCGGTTCGCTTCGACCGGGGGTGACGTCAGTGGTTCGGGCGTT 483

6231 CTACTCTGTCGATCTGTATAGTGGAGCGGAGCAGAACAGTTGAAACAGTC--- 5287

484 GGTGGGTTCTGTACCGCGCGTGTGACCGCGGAGCTGGCTATTCGATGGACCTCGCG 543

6288 -----TACGCTCCCTTAGCAGGAGAACTACTCCACCCCTGGAGCCCTCGTAGACCTAAC 6341

544 ATTTTCGGGTTTCACTTGTGGGTGCTCTCGATCATGGCGCGGATCAACATGATCAG 603

6342 ATCTTCTCTTACACCTAGCAGGTTCTCTCTATCTTAGGGCCATCAATTTTCATCACA 6401

604 ACCTTCTTGACATGCGCGCGCGGATGACGCTGCACAAAGTGGCGTTTCTCGTGG 663

6402 ACATTTATCAATAFAAAACCCCTGCCATAACCAATACCAACGCCCTTTCGTCGA 6461

664 TCGATCTTTATCAGGCTTGGGTGATCTCTGGCGGTGCGGTTCTGGGTGTCGAATC 723

6462 TCGGCTCTAATACAGAGAGTCTCTCTATCTCTCCAGTCTAGCTGTCGTCATC 6521

724 ACCATGCTGTGACCGACCGTAACTTCGGCAGCACCTTCTTCAATCTGTCGGCGCGGT 783

6522 ACTATCTACTACAGACCGCACTTCAACACACCTTCTTCAGCCCGCGGAGGAGA 6581

784 GACCGGATCTGTACCAACATCTCTGTGTTCTTTGGGACCCCGGAAGTACATCATC 843

6582 GACCCCATCTTACCAACACCTATTCGATTTTTCGGTCACCCCTGAAGTTATATCTT 6641

844 ATTCTGCGCGGCTTTGGCATCATCAGCATGTGTCGTCGACCTTCTC---GAAAAAGCG 900

6642 ATCTACAGGCTTCGAATAAATCTCCATATTTGTAACCTACTTCTCCGGAATAAAGAA 6701

901 GTCTTCGGTTACCTGCGCATGGTCTATGCAATGGTGGCAATCGGTGTTCTGGGCTTTGTC 960

Db 6702 CCAATTTGATACATAGGTATGGTCTGACGTATGATATCAATGGCTTCCTAGGGTTTATC 6761

QY 961 GTCTGGGCGCACCATGATACACCGTTGGTATGCTGCTGACCCAGCAATCTACTTCTATG 1020

Db 6762 GTGTGAGCACACCATATTTTACAGTAGGAAATAGACGTAGACACACGAGCATATTTTACC 6821

QY 1021 CTGSCCAGCATGGTATCGCGTGGCGGACCGGATTAAGATCTTCTCGTGGATCGGCACG 1080

Db 6822 TCCGCTACCAATATCATCGCTATCCCCACCGCGTCAAGATTTAGCTGACTCGCCACA 6881

QY 1081 ATGTGGGCGGCTCGGTTGAGTTCAAATCGCCGATGCTCTGGGCTTTGGCTTTATGTT 1140

Db 6882 CTCACGGAAGCAATATGAATGATCTGCTGAGTCTCTGAGCCCTAGGATTCATCTTT 6941

QY 1141 CTGTTTACCGTGGGTGGTGTGACCGGTATCGTGTGGCCCAAGCGGTCTGGACGGTGA 1200

Db 6942 CTGTTTACCGTGGGTGGTGTGACCGGTATCGTGTGGCCCAAGCGGTCTGGACGGTGA 7001

QY 1201 TATCAGACACCTATTAAGTGGTGGCGGACCTTCATTTATGATGTCGCTGGGTGGATC 1260

Db 7002 CTACAGACAGTACTAGTGTAGCCCACTTCACATATGCTTATCAATAGGAGCTGTA 7061

QY 1261 TTTCGATCTTTGCGCGGTATCTATTTTACATGCGGAAGTTCTCGGGCGCGCTTTCCCG 1320

Db 7062 TTTCGATCTTTGCGCGGTATCTATTTTACATGCGGAAGTTCTCGGGCGCGCTTTCCCG 7121

QY 1321 GAATGGGCTGCAAGCTGCACTTCTGGACCTTCTTCGCGTGGCGGAACTGACCTTCTTC 1380

Db 7122 CAACCTAGCGCAAAATCCATTTTCACTATATCATGCGCGGAAATCTTAACCTTTCTTC 7181

QY 1381 CGCAGCACTTCTTCGGGACGTGAGGTATCGCGCGCGTTTACATCGACTATCCCGAAGCC 1440

Db 7182 CCACAACACTTCTTCGGGCTATTCGGAATGCCCGAGCTTACTCGGACTACCCCGATGCA 7241

QY 1441 TTTCGCTGTGGAACAAAGTCTCTCTCTATGTTGGTTCCTGGGCTTCGCTGCTCTCTG 1500

Db 7242 TACACCATGAAACATCTCTATCTGTAGGCTCAATTCATTTCTTAACAGCAGTAATA 7301

QY 1501 TTCTTCATCTGATCTTT 1518

Db 7302 TTAATAATTTTCATGATT 7319

RESULT 14

US-10-308-264-524

; Sequence 524, Application US/10308264

; Publication No. US20040029133A1

; GENERAL INFORMATION:

; APPLICANT: Herznstadt, Corinna

; TITLE OF INVENTION: MITOCHONDRIAL DNA POLYMORPHISM

; FILE REFERENCE: 660088.461

; CURRENT APPLICATION NUMBER: US/10/308,264

; CURRENT FILING DATE: 2002-11-25

; NUMBER OF SEQ ID NOS: 697

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 524

; LENGTH: 16569

; <211> conflict

; TYPE: DNA

; ORGANISM: Homo sapiens mitochondrial

; FEATURE:

; LOCATION: 3107

; OTHER INFORMATION: "n" refers to a nucleotide at position 3107 of the

; OTHER INFORMATION: Cambridge Reference Sequence (CRS, Anderson et al.

; OTHER INFORMATION: 1981 Nature 290:457-465), later determined to be

; OTHER INFORMATION: absent from the human mtDNA sequence according to

; OTHER INFORMATION: the revised CRS (rCRS, Andrews et al. 1999 Nature

; OTHER INFORMATION: Genetics 23:147); "n" therefore represents a

; OTHER INFORMATION: single nucleotide deletion at a position corresponding to

; OTHER INFORMATION: nucleotide 3107 of SEQ ID NO:1 as provided herein.

US-10-308-264-524

Query Match

23.8%; Score 398.8; DB 16; Length 16569;

```
Best Local Similarity 58.2%; Pred. No. 5e-110;
Matches 744; Conservative 0; Mismatches 522; Indels 12; Gaps 2;

QY 244 AACGGACACTGTGGAACTGATGTTACTACCATGGTATTCATGATGTTCTTCTGATGATGTTCTTCTG 303
Db 5051 AACGACCACTATACAACTGATGTCAGCAGCCGATGATTTGTAATATCTTCTTCATA 6110
QY 304 GGTATCCCGCATTTGTCGGTGTGTTGTAATACTATGATGATGATGATGATGATGATGATGATGATG 363
Db 6111 GTAATACCCATCATATACGAGGCTTTGGCAACTGACTAGTTCCCTTAATATCCGTCGCC 6170
QY 364 CCGGATATGGCTTCCCGGATGATGAACACCTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 423
Db 6171 CCGGATATGGCTTCCCGGATGATGAACACCTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6230
QY 424 GCGATGGCGGTGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 483
Db 6231 CTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6287
QY 484 GGTGGGTGTTGTAACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543
Db 6288 -----TACCCCTCCCTTAGCAGGAACTACTCCACCCCTGAGCCCTCGTAGACCTAAC 6341
QY 544 ATTTTCGGGTTCACTTGTGCGGTGCTCTCGATCATGGGCGCATCAACATGATCAG 603
Db 6342 ATCTTCTCTTACACCTAGCAGGTGCTCTCTATCTTAGGGGCCATCAATTTCAACACA 6401
QY 604 ACCTTCTTGAACATGGCGCCCGGCATGACGTGACAAAGTGGGTGTTCTGCTGG 663
Db 6402 ACAATATCAATATAAAACCCCTGCAATACCCCAATACCAACGCCCTCTCTGCTGA 6461
QY 664 TCGATCTTTATACAGGCTTGGGTGATCTGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 723
Db 6462 TCGGCTCTTAATACAGCAGTCTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 6521
QY 724 ACATGCTCTGACCGACCGTAACTTTCGGCAGCACTTCTTCAATCTGCTGGCGCGGT 783
Db 6522 ACTATACTACTAACAGACCGCACTTCAACACACCTTCTCGACCCCGCGGAGGAGA 6581
QY 784 CACCGGATCTGTACCAACACATCTGTGTTCTTGGCACCCGAGAGTGTACATCATC 843
Db 6582 GACCCCATCTTATACCAACACATCTTCTGATTTTGGTACCCGTAAGTTTATATCTT 6641
QY 844 ATTTGCGCGGTTTGGCATCATCAGCCATGCTGCTGACCTTCTC---GAAAGAGCG 900
Db 6642 ATCTTACAGGCTTGGAAATATCTCCCATATTGTAACCTTACTCTCGGAAAAAGAA 6701
QY 901 GTCTTCGTTACTGCGGATGTTCTATGCAATGTTGGCAATCGGTGTTCTGGGCTTTGTC 960
Db 6702 CCATTTGGATACATAGGTATGTTCTGAGCTATGATCAATTTGGCTTCTAGGGTTTATC 6761
QY 961 GTCTGGCGCACCATGATGACACCGTTGTTGTTGCTGCTGACCCAGCAATCTTCTATG 1020
Db 6762 GTGTGACACACCATATTTACAGTAGGAATAGAGTAGACAGGAGCATATTTTACC 6821
QY 1021 CTGGCCACCATGTTGATCGCGTGGCGACCGGATTAAGATCTTCTGTTGGATCCGACG 1080
Db 6822 TCGCTTACCATATATCATGCTATCCCAACCGGCTCNAAGTATTTAGCTGACTGCCACA 6881
QY 1081 ATGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
Db 6882 CTCCACGAGGAAATATGAATGATCTGCTGAGTGTCTGAGCCCTAGGATTCATCTTT 6941
QY 1141 CTGTTTACCGTGGGTGTGACCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Db 6942 CTGTTTACCGTGGGTGTGACCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7001
QY 1201 TATCAGACACCTATTAAGTGTGGCGGACTTCCATTAATGATGATGCTGCTGGTGGGATC 1260
Db 7002 CTACAGACACGTAATGATGTTGAGCCCACTTCCACTATGCTCTCAATAGGAGCTGA 7061
QY 1261 TTTGCGATCTTGCGCGGTATCTACTTTTACATGCGCAAGTCTCGGCGCGGCTTCCCG 1320
```

```
Db 7062 TTTGCCATCATAGGAGGCTTCATTCACGTGATTTCCCTATTTCTAGGCTACACCTAGAC 7121
QY 1321 GAATGGGCTGCAAAAGCTGCACTTCTGACCTTCTTTCATCGGTGCGAAGCTGACGTTCTTC 1380
Db 7122 CAPACCTAGCCAAATCCATTTCACTATCATATTCATCGCGTAATCTAACTTTCTTC 7181
QY 1381 CCGGAGCACTTCTGGGACGTGAGGTATGCGGGCGGTACATCGACTATCCGGAAGCC 1440
Db 7182 CCAACACACTTCTCGGCTATCCGGAATGCCCGAGTTACTCGGACTACCCCGATGCA 7241
QY 1441 TTGCGCTGTGGAAACAAAGTCTGCTATATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Db 7242 TACACCATGAAACATCTCTATCATCTGTAGGCTCATTTCTCTAACAGAGTAATA 7301
QY 1501 TTCTTCATCTGATCTTT 1518
Db 7302 TTAATAATTTTCATGATT 7319

RESULT 15
US-10-308-264-529
; Sequence 529, Application US/10308264
; Publication No. US20040029133A1
; GENERAL INFORMATION:
; APPLICANT: HerinStadt, Corinna
; TITLE OF INVENTION: MITOCHONDRIAL DNA POLYMORPHISM
; FILE REFERENCE: 660088.461
; CURRENT APPLICATION NUMBER: US/10/308,264
; CURRENT FILING DATE: 2002-11-25
; NUMBER OF SEQ ID NOS: 697
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 529
; LENGTH: 16569
; <211> conflict
; TYPE: DNA
; ORGANISM: Homo sapiens mitochondrial
; FEATURE:
; LOCATION: 3107
; OTHER INFORMATION: "n" refers to a nucleotide at position 3107 of the
; OTHER INFORMATION: Cambridge Reference Sequence (CRS, Anderson et al.
; OTHER INFORMATION: 1981 Nature 290:457-465), later determined to be
; OTHER INFORMATION: absent from the human mtDNA sequence according to
; OTHER INFORMATION: the revised CRS (rCRS, Andrews et al. 1999 Nature
; OTHER INFORMATION: Genetics 23:147); "n" therefore represents a
; OTHER INFORMATION: single nucleotide deletion at a position corresponding to
; OTHER INFORMATION: nucleotide 3107 of SEQ ID NO:1 as provided herein.
US-10-308-264-529

Query Match 23.8%; Score 398.8; DB 16; Length 16569;
Best Local Similarity 58.2%; Pred. No. 5e-110;
Matches 744; Conservative 0; Mismatches 522; Indels 12; Gaps 2;

QY 244 AACGGACACTGTGGAACTGATGTTACTACCATGGTATTCATGATGATGTTCTTCTGATG 303
Db 6051 AACGACCACTATCAACAGCTTATCGTCAGAGCCCATGATTTGTAATATCTTCTTCATA 6110
QY 304 GGTATCCCGCATTTGTCGGTGTGTTTGGTAACCTATCTGATGCGGTGCAAAATCGCGCT 363
Db 6111 GTAATACCCATCATATACGAGGCTTTGGCAACTGACTAGTTCCCTTAATATCCGTCGCC 6170
QY 364 CCGGATATGGCTTCCCGGATGATGAACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 423
Db 6171 CCGGATATGGCTTCCCGGATGATGAACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6230
QY 424 GCGATGGCGGTGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 483
Db 6231 CTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6287
QY 484 GGTGGGTGTTGTAACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543
Db 6288 -----TACCCCTCCCTTAGCAGGAACTACTCCACCCCTGAGCCCTCGTAGACCTAAC 6341
QY 544 ATTTTCGGGTTCACTTGTGCGGTGCTCTCGATCATGGGCGCATCAACATGATCAG 603
```


Db 241 TCCTCTAGCAT-----CCTCCGGCTAGAAAGCTGGGGCTGTACAGGCTGAACCG :291
QY 488 GGGTTCTGTACCGCGCTGTCGACCGCGGAAGCTGCTATTTCGATGGACCTCGCGATT 547
Db 292 TCTACCCCACTAGCAGGCAACTGCGCCACGCGGAGCTTCGGTTGACCTAACCATCT 351
QY 548 TCGCGGTTCACTTGTTCGGGTGCTCTCGATCATGGCGCGCATCAACATGATCAGACCT 607
Db 352 TCTCCCTCCACCTGCGCGGTGTCTCATCGATTAGGGGCCAATACTTATTACCACCA 411
QY 608 TCTTGACATGCGCGCGCGCATGACGCTGCACAAAGTGGCTTCTCGTGTGCA 667
Db 412 TCATTAACATAAACCCCGAGCATCTCCCAATACAGGCCCTCTCTTCGTAAGCCG 471
QY 668 TCTTTATCAGGCTTGGCTGATCTGCTGGCGTCTGCGGTTCTGGGTGTCGAATCACCA 727
Db 472 TCTTAATTACTGCGCTCTCTCTGCTGCTGCTGCGAGTCTCGCAGCGGATCACAA 531
QY 728 TGCTGTGACGACCGTAATCTCGGACGACCTTCTTCAATCTGCTGGCGCGGTGACC 787
Db 532 TGCTTCTGACAGACCGAACTAAACACCACTTCTTCGACCTGCGGCGGAGGAC 591
QY 788 CGATTCTGTACCAACACATCTGTGGTCTTGGGACCGGAGTGTACATCATCTTC 847
Db 592 CAATTCTTACCACACTTATTCGATCTTGGCCACCCAGAGTGTACATCTCATCC 851
QY 848 TGCCCGCTTTGGCATCATCAGCGATGCTGTGACCT--TCTGAAAGCCGGTCT 904
Db 652 TACCAGGCTTTGAAATAATCTCCCATATTGCTGCTATTATGCGGCAAAAAGAACCAT 711
QY 905 TCGGTTACTGCGCATGCTATGCAATGTCGCAATCGGTCTGCGCTTTGCTGCT 964
Db 712 TCGGCTACATGGGATGTTTGGCCATGATGGTATGSCCTTCTGGGCTTCATGCT 771
QY 965 GGGCGCACCATGTACACCGTTGGTATGTCGTGACCCAGCAATCTACTTCATCTCG 1024
Db 772 GAGCCCAACATATGTTACGCTAGGAATGACGCTTGACACCGCGCTACTTCACATCC 831
QY 1025 CCACCTGTGTGTCGGTCCGACCGCATTAAGATCTCTCGTGATCGCCACGATGT 1084
Db 832 CCACATAATTAITGGCATCCCCACCGAGTAAAGTATTTAGCTAGTACGACCCCTTC 891
QY 1085 GGGCGGCTCGGTGAGTTCAATCGCGGATGCTCTGGGCTTTGCTTTATCTTCTCT 1144
Db 892 ACGCGGCTCCATTAAATAGAAACTCCAATGCTCTGAGCCCTAGGCTTCATCTTCTCT 951
QY 1145 TCACGTTGGTGGTGTACCGGTATCGTGTGGCCCAAGCGGTCTGGACCGTGATATC 1204
Db 952 TTACCGTAGCGGACTAACAGGATTTGCTTAGCCAACTCTCCCTAGACGTAGTCTCTG 1011
QY 1205 ACGACACTATTAGTGGTGGCGCACTTCATTATGATGTCGCTGGGTGCGATCTTG 1264
Db 1012 ATGACACTTACTAGTAGTTGCCCACTTCACCTAGCTCCTATCAATAGAGCGTATCG 1071
QY 1265 CGATCTTGGCGGATCTATCTTTTACATGCGGAAGTCTGGCGCGGCTTCCCGAAT 1324
Db 1072 CTATCATGGGTGCTTTTGTACACTGATTCCCACTTTTCTCAGGGTACACCTCCACAGCA 1131
QY 1325 GGGCTGAAAGCTGCACTCTCGACCTTCTTCACTGGTGGGACGTCAGTCTTCTCCGCG 1384
Db 1132 CTGGAACAAATCCATTTTGGAGTTAGTTCTGAGCGGTAACTTAACTTCTTCCCGC 1191
QY 1385 AGCACTTCTTGGGAGCTCAGGATATGCGCGCGGTTACATCGACTATCCGAGCGCTTCG 1444
Db 1192 AACACTTCTCGGACTCGCGGAATGCCCGTGCATCTCAGACTACCCAGACGCGTACA 1251
QY 1445 CGCTGTGGAACAAAGTCTGCTCTATGGTGTGGTCTGCGCTTCTGCTTCTTCT 1504
Db 1252 CGCTATGAACACAGTCTTCTCCATTTGGCTCTCTAACTCTCACTATCGCTGTCTATAT 1311
QY 1505 TCATCGTGTACTT 1517

Db 1312 TCCTATTATCCT 1324
RESULT 2
CR701206
LOCUS CR701206 1528 bp mRNA linear HTC 12-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR701206
VERSION CR701206.1 Gi:51199115
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon
ORGANISM Tetraodon
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae.
REFERENCE 1 (bases 1 to 1528)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage
: 2 rue Gaston Creneau, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : segre@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
source
1..1528
/organism="Tetraodon"
/mol_type="mRNA"
/db_xref="taxon:47144"
/tissue_type="Eyes"
ORIGIN
Query Match 24.5%; Score 409.8; DB 3; Length 1528;
Best Local Similarity 58.7%; Pred. No. 6.4e-99;
Matches 752; Conservative 0; Mismatches 517; Indels 12; Gaps 2;
QY 240 GGCAGACGACACCTGTGGAACTCATGGTACCTACATGGTATCTGATGATGTTCTT 299
Db 25 GGGTCATGACAGATTATTAATGTAATGTTACAGCCCATGCAATTGTAATATTTCTT 84
QY 300 TGTGGGTATCCCGCATGTTTCGGTGGTTTGGTAACTATCTGATGGCGGTGCAATCGG 359
Db 85 TATAGTAATGCAATCATGATTGGCGCTTTGGAACTGACTAGTGCCTCATGATTGG 144
QY 360 CGCTCCGGATATGCTTCCCGCTATGACAACTGCTGCTTCTGGCTGTTTCATTTGCCG 419
Db 145 AGCAGCTGACATGGCATTCGCCCAATTAATTAATGCTTTGGTACTACCTCTTC 204
QY 420 TACCGCATGGCGTGGCTTCGCTGTTGACCGCGGTGACGGTCAGTGGGTTTCGGG 479
Db 205 CTTCCTACTCTCTCTAGCAT-----CCTCCGGGTAGAACTGGGCTGGTACAGG 255
QY 480 GGTGGTGGGTTCTGTACCGCGCTGTGACCGCGAGCTGGCTATTTCGATGGACCT 539
Db 256 CTGAACCGTCTACCCCGCATAGCAGGCAACTGGCCCAACCGGAGCTTCGGTTGACCT 315
QY 540 CGCGATTTTCGGGTTCATCTTGTGGGTGCTCTCGATCATGGGCGCGATCAACATGAT 599
Db 316 AACCATCTTCTCCCTCCACCTGGCGGTGTCATCGATTTTAGGGCCATTAACCTTAT 375
QY 600 CAGACCTTCTTGACATGCGCGCGCGCATGACCTGACAAAGTGGCGGTGTTCTC 659
Db 376 CACCACCATCATTAACATAAAACCCCGAGCATCTCCCAATATCAGACCCCTCTCTCGT 435
QY 660 GTGTGCTATCTTATCAGGGTTCGCTGATCTCTGCTGGCGCTGCGGTTCTGGGTGCTG 719
Db 436 ATGAGCGCTCTAATTAATGCTGCGCTCTCTCTTCTTCTGCTGCTGCTGCTGCTG 495
QY 720 AATCACCATGCTGCTGACCGGCTTAATCTTCGGACGACCTTCTTCAATCTGCTGCGG 779
Db 496 GATCACAATGCTTCTTAACAGACCGGAAACCTTAACACCACTTCTTCGACCTTCAGGCGG 555

780 CGGTGACCGGATCTGTACCAACACATCGTGTGTTCTTTGGGACCCGGAAGTGTACAT 839
Db |||||
556 AGGAGACCCAAATCTTTACCAACATTAATCTTGATCTTTTGGCCACCCAGAGTGTACAT 615
Qy |||||
840 CATCATCTGCGCGCTTTGGCATCATCAGCCATCTGCTGTCGACCTTCTCG---AAAA 896
Db |||||
616 TCTCATCTTACAGGCTTTGGAATTAATCTCCCATATTTGTGCTACTATGCGCGCAAAA 675
Qy |||||
897 GCGGTCTTTCGGTTACCTGCCGATGGTCTATGCAATGGTGGCAATCGTGTCTTGGGTT 956
Db |||||
676 AGAACCATTCGGCTACATGGGCATGGTTGGGCCATGATGGCTATTGGCTTCTAGGCTT 735
Qy |||||
957 TGTGCTGCGGCGCACCATGATGACCGCTTGGTATGCTGACCGACATCTACTTT 1016
Db |||||
736 CATGCTGAGCGCCACCATATGTTACGGTAGGAATGGAAGTGTGACACCGCGCTACTT 795
Qy |||||
1017 CATGCTGGCCACCATGATCGCGGTGCCGACCGCATTAAGATCTTCTGTCGATCGC 1076
Db |||||
796 CACATCCGCCCAATATTTATGGCATCCCAACCGAGTAAAGTATTAGCTAGCTAGC 855
Qy |||||
1077 CACGATGCGGCGGCTCGGTTAGTTCAAATCGCGGATGCTCTGGGCTTTGGCTTTAT 1136
Db |||||
856 CACCCTTACGCGGCTCCATTAATAATGAGAACTCCAATGCTCTGAGCCCTAGGCTTCAT 915
Qy |||||
1137 GTTCTGTTTACCGTGGTGTGACCGGTATCGTGTGGCCCAAGCGGTCTGACCG 1196
Db |||||
916 CTTCCTTTACGTAGCGGACTAACAGCATTTGCTTAGCCAACTCTCCCTAGACGT 975
Qy |||||
1197 TGCATATACGACACTATTACGTGTGGCGCACTTCCATATGATGATGCTCGGTGTC 1256
Db |||||
976 AGTCTGTCATGACACCTACTAGTAGTGTGCCACITCCACTACGCTCTCAATAGGAGC 1035
Qy |||||
1257 GATCTTTGCACTTTCGCGGTATCTATTTTACATGCGGAAGTCTCGGCGCGCTTT 1316
Db |||||
1036 CGTATTGGCATCATGGTGTCTTTGTACATGATTCACATTTCTCAGGATACACCT 1095
Qy |||||
1317 CCGGAATGGCTGCAAACTGCACTTCTGGACCTTCTTCATCGTGGCAAGCTCACGTT 1376
Db |||||
1096 CCACAGCACTGAAACAAAATCCATTTGGAGTTATTTGCTAGGCGTTAACCTTAACCT 1155
Qy |||||
1377 CTTCCGAGCACTTCTGGAGGTGAGGTATGCGCGCTTACATCGATATCCGATCCCGA 1436
Db |||||
1156 CTTCCGCAACACTTCTCGGACTCGCGGAATGCGCGCTGATACACTAGATACCCAGA 1215
Qy |||||
1437 AGCTTTCGGCTGTGAAACAAAGTCTGCTCTATGCTGCTTCTCGGCTTCTCGCTCGTT 1496
Db |||||
1216 CGGTTACAGCTATGAAACACAGTTTCTCCATTTGCTCTCTAACTCATTACGCTGT 1275
Qy |||||
1497 CTTGTTCTTCATCGTGAATCTT 1517
Db |||||
1276 CATTATATTCCTATTATCT 1296

RESULT 3

CR694572
LOCUS 1614 bp mRNA linear HTC 12-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR694572
VERSION CR694572.1 GI:51192479
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon
ORGANISM Tetraodon

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae.

1 (bases 1 to 1614)

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
(2 rue Gaston Crémieux, CP 5706 - 91057 EVRY cedex - FRANCE)
(E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT

The sequences are based on single pass reads.

More information available at
<http://www.genoscope.cns.fr/tetraodon>.

FEATURES

source

1..1614
/organism="Tetraodon"
/mol_type="mRNA"
/db_xref="taxon:47144"
/tissue_type="Eyes"

ORIGIN

Query Match 24.4%; Score 408.2; DB 3; Length 1614;
Best Local Similarity 58.6%; Pred. No. 1.8e-98;
Matches 751; Conservative 0; Mismatches 518; Indels 12; Gaps 2;

Qy 240 GCGGAACGACACCTGTGGAAAGTCATGTTACCTACCATGTTATCTGATGATGTTCTT 299
Db |||||
135 GGGTGATGACAGATTTATATGTAATCGTTACAGCCACGCATTTGTAATAATTTCTT 194
Qy |||||
300 TGTGGGTATCCCGCATTTGTCGGTGTGTTTGGTAACTATCTGATGCGCTGCAAAATCGG 359
Db |||||
195 TATAGTAATGCAATCAATGATGGCGCTTTGAAACTGACTAGTCCCTCATGATGG 254
Qy |||||
360 CGTTCGGATATGCGCTTCCCGGTATGAACAACTGTCGTTCTGGGTGTTCAATGCGGG 419
Db |||||
255 AGACCTGACATGGCATTTCCCGGATAAATAACATAAGCTTTTGGTTACTACCTCCTTC 314
Qy |||||
420 TACCGCATGGCGTGGCTTCGCTGTTTCGCACCGGCGGTGACGTCAGCTGGTTCGGG 479
Db |||||
315 CTTCTACTCTCTAGCAT-----CTCGGGGTAGAAGCTGGGCTGGTACAGG 365
Qy |||||
480 CGTTGGTTGGGTTCTGTACCCGCGCTGTGCAACCGCGAAGCTGGCTATTTCGATGACCT 539
Db |||||
366 CTGAACCGTCTACCCCCCACTAGCAGGCAACCTGGCCACCGCGAGCTTCGGTTGACCT 425
Qy |||||
540 CGCATTTTCGGGTTCACTTGTGGGTGCTCTCGATCATGGCGCGATCAACATGAT 599
Db |||||
426 AACCATCTTCTCCCTCCACCTGGCGGTGCTCATCGATTTTAGGGCCATTAACCTTAT 485
Qy |||||
600 CACGACCTTCTGAACATGGCGCCCGCGCATGACGCTGCACAAAGTCGCGTTGTTCTC 659
Db |||||
486 TACACACATCATTAACATAAACCACGCACTCCCAATATCAGACCCCTCTCTCGT 545
Qy |||||
660 GTGTCGATCTTTATACAGGTTGGCTGATCTCTGTCGCGCTGCCGTTCTGGCTGTGC 719
Db |||||
546 ATGAGCGCTCTTAATTAAGTCTGCTCTCTCTGCTCTCTGCTCTCTGCTCTGCGAGCGG 605
Qy |||||
720 AATCACCATGCTGTGACCGACCGTAACCTTCGGCAGCACCTTCTTCAATCTGTCGCGG 779
Db |||||
606 GATCACATGCTCTGACAGACCGAACCCTAACACACCACTTCTTCGACCTTCGAGCGG 665
Qy |||||
780 CGGTGACCCGATCTGTACCAACACATCTGTGTTCTTTGGGCAACCGGAAGTGTACAT 839
Db |||||
666 AGGGAGCCCAATCTTTACCAACACTTATTCTGATTCTTTGGCCACCCAGAGTGTACAT 725
Qy |||||
840 CATCATTTGCGCGCTTTGGCATCATGACCATGTCGTGTCGACCTTCTC---GAAAAA 896
Db |||||
726 TCTCATCTACAGGCTTTGGAATAATCTCCCATTTGTTGCTTATTATGCGCGCAAAA 785
Qy |||||
897 GCGGCTCTTCGGTTACCTGCGATGGTCTATGCAATGGTGGCAATCGGTGTTCTTGGGCTT 956
Db |||||
786 AGAACCATTCGGCTACATGGGCATGGTTGAGCCATCATGGCTATTGGCTTCTGGGCTT 845
Qy |||||
957 TGTGCTGTGGCGGACACATGATGACCGTTGGTATGCTGCTGACCGCAATCTTACTT 1016
Db |||||
846 CATGCTCTGAGCCACCATATGTTTACGGTAGGAATGAGCTTGACACCCGCGCTACTT 905
Qy |||||
1017 CATGCTGGCCACCATGTTGATGCGGCTGCGGACCGCATTAAGATCTTCTGTCGATCGC 1076
Db |||||
906 CACATCCGCCACATAATATTGGCATCCCCACCGAGTAAAGTATTAGCTACATAGC 965
Qy |||||
1077 CACGATGTGGCGGCTCGGTTGAGTTCAAAATCGCGCATGCTCTGGGCTTTTGGCTTTAT 1136
Db |||||
966 CACCTTTACGCGGCTCCATTAATAAGAAACTCCAATGCTCTGAGCCCTAGGCTTCAT 1025

Db 1332 CGGTACACGCTATGAACACAGTTTCTCCATTGGCTCTCTAATCTCACTTATCGCTGT 1391
QY 1497 CCTCTTCTTCATCGTGATCTT 1517
Db 1392 CATTATATCTTATTTATCT 1412

RESULT 5
CR704060 1638 bp mRNA linear HTC 12-AUG-2004
LOCUS Tetraodon nigroviridis full-length cDNA.
DEFINITION CR704060
ACCESSION CR704060
VERSION CR704060.1 GI:51201969
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon
ORGANISM Tetraodon
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae.
1 (bases 1 to 1638)
Genoscope.
Direct Submission
Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
source
1..1638
/organism="Tetraodon"
/mol_type="mRNA"
/db_xref="taxon:47144"
/tissue_type="Eyes"

ORIGIN
Query Match 24.4%; Score 408.2; DB 3; Length 1638;
Best Local Similarity 58.6%; Pred. No. 1.8e-98;
Matches 751; Conservative 0; Mismatches 518; Indels 12; Gaps 2;
QY 240 GGGGAACGACACCTGTGAAAGTCATGTTGATTTACCTACCACTGATCTGATGATCTTCT 299
Db 141 GGGTGATGACCAAGATTATATGTAATCTGTACGCCACGATTTGTTAATATTTTCTT 200
QY 300 TGTGGGTATCCCGCATTTGTTCGGTGGTTTGTGTAATCTGATGCCCTCCAAATCGG 359
Db 201 TATAGTAATGCCAATCATGATTGGCGCTTTGGAACTGACTAGTGCCTCATGATTGG 260
QY 360 CGTCCGGATATGGCTTCCCGGATGAACAACTGTCTGCTGCTGTTCAATTGCCGG 419
Db 261 AGCACCTGACATGGCATTCGCCGAATAAATAACATAAGCTTTTGGTTACTACCTCCTTC 320
QY 420 TACCGGATGGGCTGGCTTCGCTGTTCGACCGGCGGTGACGGTCACTGGGTTCGGG 479
Db 321 CTTCTACTCCTCTAGCAT-----CCTCGGGGTAGAAGCTGGGGCTGTACAGG 371
QY 480 CGTTGGTTGGGTCTCTACCCGCGCTGTGACCCCGGAAGCTGGCTATCGATGACCT 539
Db 372 CTGACCGCTACCCGCCACTAGCAGCAACCTGGCCACCGCGAGCTTCGTTGACCT 431
QY 540 CGGATTTTCGGGTTTACCTTGTGGGTGCTCTCTGATCATGGCGCGATCAACATGAT 599
Db 432 AACCATCTTCTCCCTCCACCTGGCGGTGTCTATCGATTTTAGGGGCCAATTAATTTAT 491
QY 600 CACGACCTCTTTGAACATGGCGGCCCGGCATGACGCTGCACAAAGTGGCGTTGTCTC 659
Db 492 TACACCATATTAACATAAACCACGCACTCCCAATATACAGACCTCTCTTCGT 551
QY 660 GTGTCGATCTTATACAGGCTTGGCTGATCTTCGTGGCGCTGCCGTTCTGCTGGTGC 719
Db 552 ATGACCGCTCTTAATTAATGCTGCGCTCTCTCTCCGCTCTCTGCGCTCTCTCGACCGG 611

QY 720 AATCACATGCTGCTGACCGACCGTAACCTTCGGACGACCTTCTTCAATCTCTGCTGGGG 779
Db 612 GATCACAATGCTTCTGACAGACCGAAACCTAAACACCACTTCTTTCGACCTTCGAGCGG 671
QY 780 CGGTGACCGGATTTGTACCAACACATCTCTGTGGTCTTGGGACCGCGGAAGTGTACAT 839
Db 672 AGGGGACCAATTTCTTTACCAACACTTATCTGATCTTTTGGCCACCCAGAGTGTACAT 731
QY 840 CATCATCTTCGCCGGCTTTGGCATCATCAGCCATGTCGTGTCGACCTTCTC--GA AAAA 896
Db 732 TCTCATCTTACAGGCTTTGGATATCTCCCATATTTGTCCTATTATGCGGCAAAAA 791
QY 897 GCGGCTTTCGGTTTACTGCCGATGCTATGCAATGTTGCAATGGTGGCAATCGGTGTTCTGGCTT 956
Db 792 AGAACCAATTCGGCTACATGGGATGCTTTGAGCCATGATGGCTATTGGGCTTCTGGGCTT 851
QY 957 TGTGCTGTCGGGCGACCATGATACACGCTTGTGATGTGCTGACCGACCAATCTTACTT 1016
Db 852 CATCTGCTGAGCCACCATATGTTTACGTTAGGAATGAGAGTTGACACCGCGCTTACTT 911
QY 1017 CATGCTGGCCACCATGCTGATCGCGTGGCGGACCGGCAATTAAGATCTTCTGTTGATCGC 1076
Db 912 CACATCGGCCACCAATAATTATTGCCATCCCGAGTAAAGATTTAGCTGACTAGC 971
QY 1077 CACGATGTCGGGCGGCTCGGTTGAGTTCAATCGCGATGCTCTGGGCTTTTGGCTTTAT 1136
Db 972 CACCTTTCACGGCGGCTTCAATTAATGAGAACTCCAATGCTCTGAGCCCTAGGCTTCAT 1031
QY 1137 GTTCTCTTTCACCGTGGGCTGTGACCGGTATGCTGTGGCCCAAGCGGGTCTGGACCG 1196
Db 1032 CTTCTCTTTTACCGTAGGCGGACTTAACAGGCAATTGCTTAGCCAACTCTCTCCCTAGACGT 1091
QY 1197 TGCATATCAGACACCTATTACGTTGGGCGGCTTCCATATGATGATCGGTGGGTGC 1256
Db 1092 AGTCTGCTATGACACTTACTACGTAGTTGCCACTTCCCACTGCTCTCAATAGAGC 1151
QY 1257 GATCTTTGCGATCTTCGCGGTATCTACTTTTACATGCGGAAGTTCTCGGCGCGGCTT 1316
Db 1152 GATTTGCTATCATGGGCTTTTGTACACTGATTTCCACTTTTCTCAGGTACACCT 1211
QY 1317 CCGGAATGGGCTCAAGCTGCACTTCTGACCTTCTTCTCATCGTGGGAACGTCAGCTT 1376
Db 1212 CCACGACACCTTGAACAAAAATCCATTTTGGAGTTATGTTCTGAGCGCTTAACCTAACCTT 1271
QY 1377 CTTCCCGCAGACCTTCTGGGACGTCAGGATGTCGCGCGCTTACATCGACTATCCCGA 1436
Db 1272 CTTCCCGCAACACTTCTCGGACTCGCGGGAATGCCCGCTCGATCTCAGACTACCCAGA 1331
QY 1437 AGCTTTCGCGCTGTGGAACAAAGTCTCTGCTATGTTGGTTCCTGGCTTCGCTCGCT 1496
Db 1332 CGGTACACGCTATGAAACACAGTTTCTCATTGGCTCTCTAATCTCACTTATCGCTGT 1391
QY 1497 CTTGTTCTTCTGATGATCTT 1517
Db 1392 CATTATATTCCTATTATTCCT 1412

RESULT 6
CR691904 1642 bp mRNA linear HTC 12-AUG-2004
LOCUS Tetraodon nigroviridis full-length cDNA.
DEFINITION CR691904
ACCESSION CR691904
VERSION CR691904.1 GI:51189811
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon
ORGANISM Tetraodon
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae.
1 (bases 1 to 1642)
Genoscope.
Direct Submission


```

182 TATAGTAATGCAATCATGATGTCGGCGCTTGGAACTGACTAGTGCCTCATGATGG 241
360 CGCTCCGGATATGGCTTCCCGGTATGAACAACCTGTCTGCTTCTGGGTGTCTATGCGG 419
242 AGCACTGACATGCAATCCCGGATGAATAAATAAATAAAGCTTTTGGTTACTACCTCCTC 301
420 TACCGGATGGCGTGGCTTGGCTTGGCTGTCGACCGGGGGGTGACGGTCACTGGGTGGG 479
302 CTTCCTACTCTCTCTAGCAT-----CTCCGGCGTGAAGCTGGGGTGGTACAGG 352
480 CGTTGGTTGGGTCTGTACCGCGCGCTGTGACACCGGGAAGCTGGCTATTTCGATGGACCT 539
353 CTGAACCGTCTACCCCCACTAGCAGGCAACCTGGCCACACCGCGAGCTTCGGTTGACCT 412
540 CGCGATTTTCGGGTACTTGTTCGGGTGCTCTCTGATCATGGCGGGGATCAACATGAT 599
413 AACCATCTCTCCCTCCATCTGGCGGTGTCTCATCGATTTTAGGGGCCATTAACCTTAT 472
600 CACGACCTTCTTGACATGCGCGCCCGGATGACGCTGCACAAAGTGCCTTGTCTC 659
473 TACCAACCATTAATAAATAAACCACGCAATCTCCCAATACAGACCTCTCTTCGT 532
660 GTGCTCGATCTTTATACAGGCTTGGCTGATCTGCTGGCGTGGCGGTCTGGCTGGTC 719
533 ATGAGCGCTCTTAATTAAGTCTGCTCTCTGCTGTCTGCAAGTCTCGAGCGG 592
720 AATCAGATGCTGTGACCGACCGTAACCTCGGACGACCTTCTCAATCTGCTGGCG 779
593 GATCAATATGCTTGTGACAGCCGGAACCTAAGACCACTTCTTCGACCTTGAGCGG 652
780 CGGTACCGGCTTCTGTACCAACACATCTGTGTGTTCTTGGGCACCGGAAGTGTACAT 839
653 AGGAGACCAATCTTTACCAACATTTCTGTGATCTTTGGCCACCCAGAGTGTACAT 712
840 CATATTCGCGCGCTTGGCATCATACGACGATGCTGTGCGACCTCTC---GAAAA 896
713 TCTCATCTTACAGGCTTGGAAATACTCCCATATTTGTCCTATTATGCGGCAAAA 772
897 GCGGCTCTTCGCTTACCTGCCATGCTATGCAATGGTGGCAATCGGTGTCTGGGCT 956
773 AGAACCAITTCGGCTATCGGGCATGTTTGAAGCATGATGGCTATTTGGCTTCTGGT 832
957 TGTCTGCTGGGGGCGCACCATATGACCGTTCGGTATGCTGACCCAGCAATCCTACT 1016
833 CATGCTGTCAGCCCAACCATATGTTACGGTGAAGATGACGTTGACACCGCGCTACT 892
1017 CATGTCGGCCACCATGCTGATCGCGTGGCGGACCGCATTAAGATCTTCTGCTGATCG 1076
893 AACATCGGCCACAATAATTAATGCCATCCCGGAGTAAAGTATTTAGCTGATGAC 952
1077 CACGATGTCGGGCGGCTCGGTTGAGTTCAAATGCGCGATGCTCTGGGCTTTGGCTTAT 1136
953 CACCTTCACGGGGCTCCCATTAATGAGAACTCCATGCTCTGAGCCTTAGGCTTCAT 1012
1137 GTTCCTGTTACCGTGGGTGTGACCGGTATCGTGTGGCCCAAGCGGTTGACCG 1196
1013 CTTCTCTTTTACCGTAGCGGACTAAGGCAATTTGCTTAGGCACTCTCCCTAGAGCT 1072
1197 TGCATATCAGACACCTATTACGTTGGTGGCGACTTCCATATTGATGTGCTGGGTGC 1256
1073 AGTCTGCTAGACATTTACTAGTAGTTGCCACTTCCACTAGCTCTCATCAATAGGAG 1132
1257 GATCTTTGGCATCTTGGCGGTATCTATCTTTTACATGCGGAAGTCTTCGGGCGGCTTT 1316
1133 CGTATTGCTATCATGGGTGCTTTGTACACTGATTCCCACTTTTCTCAGGGTACACCT 1192
1317 CCGGAATGGGCTGCAAGCTGACATCTGACCTTCTTCATCGGTGCGAACTGACGTT 1376
1193 CCACAGACCTGAAACAAAAATCCATTTTGAGTTATGTTGAGGCTTTAACTTAACCT 1252
1377 CTTCCCGCAGCACTTCTCGGAGGCTCAGGGTATGCGCGCGGTTACATGACTATCCGA 1436
```

```

1253 CTTCCCCCAACACTTCTCGGACTCGCGGAATGCCCGTGCATACTCAGACTACCCAGA 1312
1437 AGCTTCGGCGTGTGGACAAAGTCTCGTCTATGCTGCTGCTTCCGCTTCGCTCGTT 1496
1313 CCGGTACAGCTATGAACACAGAGTTTCCCAATGGCTCTCTAAATCTCACTATCGCT 1372
1497 CCGTGTCTTTCATCGTATCTT 1517
1373 CATTATATTCCTATTATCTT 1393

CR695115      1590 bp      mRNA      linear      HTC 12-AUG-2004
Tetraodon nigroviridis full-length cDNA.
CR695115
ACCESSION
CR695115.1 GI:51193022
VERSION
HTC; cDNA; full-length; Tetraodon nigroviridis.
KEYWORDS
Tetraodon
SOURCE
ORGANISM
Tetraodon
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae.
REFERENCE
1. (bases 1 to 1590)
Genoscope.
AUTHORS
Direct Submission
TITLE
Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
JOURNAL
: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(H-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT
The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
Location/Qualifiers
1..1590
/organism="Tetraodon"
/mol_type="mRNA"
/db_xref="taxon:47144"
/tissue_type="Eyes"
ORIGIN
Query Match 24.1%; Score 403.2; DB 3; Length 1590;
Best Local Similarity 58.5%; Pred. No. 3.9e-97;
Matches 746; Conservative 0; Mismatches 518; Indels 12; Gaps 2;

QY 245 ACGGACACCTGTGGAACTGATGTTACCTACCTACCTACCTACCTACCTACCTACCT 304
DB 115 ATGACCAATTTATAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 174
QY 305 GTATCCCGCATTTCTGCGTGGTTTGGTAACTATCTGATGCGGCTGCAAAATCGGCGTC 364
DB 175 TAATGCCAATCATGATTTGGCGGCTTTGGAACTGACTAGTGCCCTCATGATTGGAGCAC 234
QY 365 CGGATATGGCTTCCCGGTATGAACAACTGCTGCTTCTGCTGCTTCTGCTGCTTCTGCTGCT 424
DB 235 CTGACATGGCATTCCTCCGGAATAAATAACATAAGCTTTTGGTTACTACCTCTCTCTCTCC 294
QY 425 CGATGGGCGTTCGCTGTTGCGACCGGGCGGTGACGGTCACTGGGTTCGGGCGGTG 484
DB 295 TACTCTCTTAGCAT-----CTTCGGCGTAGAAGCTGGGGTGGTACAGGCTGAA 345
QY 485 GTTGGGTTCTGTACCGCGCGGTGTGACCCCGCAAGTGGGTATTCGATGAGACCTCGGGA 544
DB 346 CCGTCTACCCCACTAGCAGGCAACCTGGGCCACCGCGGAGCTTCGTTGACCTAACCA 405
QY 545 TTTTCGGGTTCACTTGTGCGGTGCTCTCGATCATGGGCGCGATCAACATGATCAGGA 604
DB 406 TCTTCTCCCTCCATCTGGCGGTGTCTCATCGATTTTAGGGGCCATTAACCTTATTACCA 465
QY 605 CTTCTTGAACATGCGCCCGCGCATGACGCTGCACAAAGTCCGTTGTTCTCTGCTG 664
DB 466 CCATCATTTAAATAAAACCCCGAGCCATCTCCCAATATCAGACCCCTCTCTCTGATGAG 525
```

QY 665 CGATCTTTATCACGGTTGGCTGATCTCTGTGGCGCTGCCGGTCTGGCTGTGGCTGTGCAATCA 724
 Db 526 CCGTCTTAATTAATCTAGTCCGCTCTCTCTGCTGTCTTGGCAGTACTCGCAGCCGGGATCA 585
 QY 725 CCATGCTGTGACGACCGTAACCTTCGCGACGACCTCTCTCAATCTCTGCTGGCGGGGTG 784
 Db 586 CAATGCTTCTGACAGACGAAACCTTAACACACTTCTCTGACCTTGCAGGCGGAGGGG 645
 QY 785 ACCGATCTGTACCAACACATCTGTGTCTTTGGGACCCGGAAGTGTACATCATCA 844
 Db 646 ACCCAATCTTTACCAACACTTATCTTGATCTTTGGCCACCCAGAGTGTACATCTCA 705
 QY 845 TTCTGCCCGGCTTTGGCATCATGACCATCTGCTGTGACCT---TCTGAAAAGCCCGG 901
 Db 706 TCTTACAGGCTTTGGAATTAATCTCCATATTGTGCTTAATGCGGCAAAAGAAC 765
 QY 902 TCTTCGGTTACCTGCCGATGCTATGCAATGTGGCAATCGGTGTTCTGGGCTTTGTG 961
 Db 766 CATTCGGCTACATGGGCATGTTTGACCATGATGGCTATGACCTTCTGGGTTTCATCG 825
 QY 962 TCTGGGCGCACCATGATACCGTTGGTATGCTGACCCAGCAATCTTACTTTCATGC 1021
 Db 826 TCTGAGCCCAACATATGTTTACGGTAGGAATGGACCTTGACACCCGCGCTACTTCCAT 885
 QY 1022 TGGCCACCATGTGTATCGCGTCCGACCGCATTAAGATCTTCTGCTGATCGCCACGA 1081
 Db 886 CCGCCACAATAATTATGCCATCCCAACCGAGTAAGATTTAGCTGACTAGCCACCC 945
 QY 1082 TGTGGGCGGCTCGTTGAGTTCAATGCGCGATGCTCTGGGCTTTGGCTTTATGTTCC 1141
 Db 946 TTCACGCGGCTCCATTAATAGAAACTCCAAATGCTCTGAGCCCTAGGCTTTCATCTCC 1005
 QY 1142 TGTTCACGTTGGTGTGACCGGTATCGTGTGGCCCAAGCGGTCTGGACCGTGCAT 1201
 Db 1006 TCTTTACCGTAGCGGACTAACAGGATGCTCTAGCCCAACTCTCCCTAGACGTAGTCC 1065
 QY 1202 ATCAGACACTATTACGTGTGGCGCACTTCCATTATGATGTCGTGGTGGGATCT 1261
 Db 1066 TGCATGAACCTTACTAGTAGTTGGCCACTTCACATACCTCTATCAATAGAGCGGTAT 1125
 QY 1262 TTGCGATCTTCGCGGTAATCTACTTTTACATGCGGAATTTCTGGCGCGGCTTTCCCGG 1321
 Db 1126 TCCTATATCGGTGCTTTGTACACTGATTCCTCACTTTCTCAGGCTACACCTCCACA 1185
 QY 1322 AATGGGCTGCAAGCTGCACTTCTGACCTTCTCATCGTGGCAACGTCAGTCTTCTCC 1381
 Db 1186 GCACCTGAACAAAATCCATTTTGGAGTTATGTTCTAGGCGTTAACTAACTTCTTCC 1245
 QY 1382 CGCAGCACTTCTGGGACGTAGGTATGCGGCGGCTTTACATGCACTATCCCGAAGCT 1441
 Db 1246 CCAAAACTTCTCGGACTCGCGGAATCCCGCTCGATPACTCAGACTACCCAGACGGT 1305
 QY 1442 TCAGCTGTGGAAAGCTCGTCTATGTCGTCTCTGCGCTTCTGCGCTTCTGCTGTCT 1501
 Db 1306 ACACGCTATGAACACAGTTTCTCAATGGCTCTCAATCTCACTTATCGCTGTCAITA 1365
 QY 1502 TCTTCATCGTATCTT 1517
 Db 1366 TATTCCTATTATCTCT 1381

RESULT 9

CR689102
 LOCUS Tetraodon nigroviridis full-length cDNA. linear HTC 12-AUG-2004
 DEFINITION Tetraodon nigroviridis full-length cDNA.
 ACCESSION CR689102
 VERSION CR689102.1 GI:51187009
 KEYWORDS HTC: cDNA; full-length; Tetraodon nigroviridis.
 SOURCE Tetraodon
 ORGANISM Tetraodon
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes;

Tetraodontidae; Tetraodontidae.

1 (bases 1 to 1426)

Genoscope.

Direct Submission

Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage

: 2 rue Gaston Cremlieux, CP 5706 - 91057 EVRY cedex - FRANCE

(E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

The sequences are based on single pass reads.

More information available at

http://www.genoscope.cns.fr/tetraodon.

Location/Qualifiers

1..1426

/organism="Tetraodon"

/mol_type="mRNA"

/db_xref="taxon:47144"

/tissue_type="Eyes"

1061 TCCTCTCGTGATGCCACGATGGGGCGGCTCGGTGAGTTCAAAATCGCGATGCTCT 1120
 Db
 774 TATTAGCTGACTAGCCACCCCTTACGGGGCTCCATTAATGAGAACTCCAAATGCTCT 833
 Qy 1121 GGGGCTTTGGCTTATGTTCTGTTTACCGTGGGTGGTGTGACCGGTATGCTGTGCCCC 1180
 Db 834 GAGCCCTAGGCTTCACTTCTCTTTACCGTAGCGGACTAACAGGCAATGTTCTTAGCCA 893
 Qy 1181 AAGCGGGTCTGGACGCTGATACACGACACTTATTAGTGGTGGCGCACTTCCATTATG 1240
 Db 894 ACTCTCCCTAGAGTAGTCTGATGACACTACTAGTAGTTGCCCACTTCCACTAG 953
 Qy 1241 TGATGTCGCTGGGTGCGAATTTTGGATCTTTCGGCGGTATCTATTTTACATGCGGAAGT 1300
 Db 954 TCCATCAATAGGACCGCTATTTCGCAATCATGGGTGCTTTTGTACACTGATTCACCTT 1013
 Qy 1301 TCTGGGCGGCTTTCGGGAATGGGCTGCAAGCTGCACTTCTGGACCTTCTTTCATCG 1360
 Db 1014 TCTAGGATACACCCCTCCAGACCTGAAACAAATCCATTTTGGAGTATGTTCTAG 1073
 Qy 1361 GTGCGAAGCTCACGTTTCCCGCAGCACTTCTTGGACGTCGAGGTATGCGGCGGCTT 1420
 Db 1074 GGTTTAACCCTAACCTTCTCCCAACACTTCTCGGACTCGCGGAATGCGCGTGGAT 1133
 Qy 1421 ACATGACTATCCGCAAGCTTCCGGCTGTGGACAAAGTCTCTCTATGTTGGTTC 1480
 Db 1134 ACTCAGATTACCGACGCTGACAGCTATGAACACAGTTTCTCCATTGGCTCTCTAA 1193
 Qy 1481 TGGCCTTCGCTCGTTCCTGTTCTTTCATCGTATCTT 1517
 Db 1194 TCTCACTATCGCTGCTATATATTCCTATTATCTT 1230

RESULT 10
 BC024102
 LOCUS
 DEFINITION Homo sapiens, clone IMAGE:3508235, mRNA.
 ACCESSION BC024102
 VERSION BC024102.1 GI:18848186
 KEYWORDS
 SOURCE HTC.
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 3167)
 Strausberg,R.
 Direct Submission
 Submitted (19-FEB-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
 Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
 Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
 Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott,
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
 George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 23 Row: e Column: 21

This clone has the following problem: frame shifted.

FEATURES

Location/Qualifiers
 1..3167
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3508235"
 /tissue_type="Brain, neuroblastoma"
 /clone_lib="NIH MGC_19"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"

ORIGIN

Query Match 23.7%; Score 397.2; DB 3; Length 3167;
 Best Local Similarity 58.1%; Pred. No. 1.9e-95; Indels 12; Gaps 2;
 Matches 743; Conservative 0; Mismatches 523;

Qy 244 AACGGACACCTGGGAACGTCATGGTTACCTACCATGATTTCTGATGATGTTCTTTGTG 303
 Db 1697 AAGGACCACATCTACAACGTTATCGTCACAGCCCATGCAATTTGTAATAATCTCTTCA 1756
 Qy 304 GGTATCCCGCATGTTCCGTGGTGTGTTGTAATATCTGATGCCGTGCAAAATCGCGGT 363
 Db 1757 GTAATACCCATCAATAATCGAGGCTTTGGCACTGACTAGTTCCCTTAATAATTCGTTG 1916
 Qy 364 CCGGATATGGCCTTCCGGGTATGAACAACTGCTCTGCTTCTGGCTGTTCTATTCGCGTACC 423
 Db 1817 CCGGATATGGCTTCCCGCATAAACATTAAGCTTCTGACTCTTACCTCCTCTCTC 1876
 Qy 424 GCGATGGGCTGGCTTCCGCTGTTCCGACCGGGGGGTGACGGTCAGCTGGTGGGCGGT 483
 Db 1877 CTACTCCTGCTGCATCTGCTATAGTGGAGCGCGGAGCAGGAACAGGTTGAACAGTC --- 1933
 Qy 484 GGTGGGTTCTGTACCGCGCTGTGACCGCGGAAGCTGGCTATTTCGATGGACCTCGCG 543
 Db 1934 -----TACCTCCCTTAGCAGGAACTACTCCACCTTGGAGCTCCGTAGACCTAAC 1987
 Qy 544 ATTTTCGGCTTCACTTGTGGGTGCTCTCTGATCATGGGCGCGATCAACATGATCAAG 603
 Db 1988 ATCTTCTCCTTACACCTAGCAGGTGCTCTCTCTATCTTAGGGGCCATCAATTTCA 2047
 Qy 604 ACCTTCTTGACATGCGCGCGCGCATGACGCTGCACAAAGTGGCTGTTCTCGTG 663
 Db 2048 ACAATTATCAATAAAACCCCTGGCATAAACCAATACCAAGCCCTCTTCGTTCTGA 2107
 Qy 664 TCGATCTTTATCAGCGCTTGGCTGATCTCTGCTGGCGTGGCGGTTCTGGTGGTCAATC 723
 Db 2108 TCGTCTTATCAGCAGTCTCTACTTCTCTCTCTCCAGTCTCTAGCTGTGGCATC 2167
 Qy 724 ACAATGCTGTGACCGGACCTTAATCTTGGGACGACCTTCTTCAATCTGTGGCGGCGT 783
 Db 2168 ACTATACTACTAAGACCGCAACCTCAACACCACCTTCTTCGACCCCGCGGAGGGA 2227
 Qy 784 GACCCGATTCTGTACCAACACATCTCTGTTCTTTGGGACCCCGGAAGTGTACATCATC 843
 Db 2228 GACCCCAATCTATACCAACACCTTATCTGATTTTCGGTCACTCTGAATTTATATCTT 2287
 Qy 844 ATTTCCCGCGCTTGGCATCATCAGCCATGTCGTGTGACCTTCTC ---GAAAGCCG 900
 Db 2288 ATCTTACAGGCTTCGGAATAATCTCCCAATTTGTAATTTACTACTCCGGAAGAAAGAA 2347
 Qy 901 GTCTTCGGTTACCTGCGGATGTTATGCAATGGTGGCAATCGGTGTTCTGGGCTTTGTC 960
 Db 2348 CCATTTGGATACATAGGTATGGTCTGAGCTATGATATCAATTTGGCTCTTAGGTTTATC 2407
 Qy 961 GTCTGGGCGGACCATGATGACCGTTGGTATGTCGTGACCCAGCAATCCTACTTCATG 1020
 Db 2408 GTGTGACACACCATATATTACAGTAGGAATAGACGTAGACACACAGCATATTTTACC 2467
 Qy 1021 CTGGCCACCATGTTGATCGCGGTGCCGCGGATTAAGATCTTCTCTGTTGATCGCCACG 1080
 Db 2468 TCCGCTACCAATATCATCGCTATCCCCACCGCGCTCAAGATTATTAGTACTGCTGCCACA 2527


```

Db      1151 ATTGGCATATAGAGGCTTCATTCACTGATTTCCCTATTTCTCAGGCTACACCTAGA 1210
Qy      1320 GGAATGGCTGCAAGCTGACTTCTGGACCTTCTTTCATCGGTGGGAACGTCAGTTCCT 1379
Db      1211 CCAAACTACGCCAAATCCATTTTCGTTATCATATTCATCGCGCTAAATCTAACTTCTT 1270
Qy      1380 CCGGAGCACTTCTTGGAGCTGACGGGTATCCGCGCGGTTTACATCGACTATCCCGAAGC 1439
Db      1271 CCACAAACATTTCTCGGCTTATCGGAATGCCCGAGTTTACTCGGACTATCCGATGC 1330
Qy      1440 CTTCCGCTGTGGAAACAAAGTCTGCTCTATGTTGGTCTGCTTCCGCTTCGCTGTTCT 1499
Db      1331 ATACACCAATGAATATCTATCATCTGTAGGCTCATTCATTTCTCTAACAGCAGTAAT 1390
Qy      1500 GTTCTTCATCGTATCTTT 1518
Db      1391 ATTAATAATTTTCATAAT 1409

RESULT 12
LOCUS   CR590798
DEFINITION 1499 bp mRNA linear HTC 21-JUL-2004
full-length cDNA clone CS0DC025YC06 of Neuroblastoma Cot
25-normalized of Homo sapiens (human).
ACCESSION CR590798
VERSION 1
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Faraday Avenue
2 (bases 1 to 1499)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
Bp 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NoI-oligo(dm) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
1..1499
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC025YC06"
/tissue_type="Neuroblastoma Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 23.7%; Score 396.2; DB 3; Length 1499;
Best Local Similarity 58.1%; Pred.No. 2.9e-95;
Matches 743; Conservative 0; Mismatches 523; Indels 13; Gaps 2;

Qy      244 AAGGACACTGTGGAGCTGATGTTACTACCTAGTATTCGATGATGTTCTTTTG 303
Db      141 AAGACACATCTACAACTGTTTCGTCAGCCCATGATTTGTAATAATCTTTCATA 200
Qy      304 GGATCCCGCAATGTTTGGTGGTTTGGTAATATCTGATCCGCTGCAATTCGGCGCT 363
Db      201 GTAATACCATATAATCGGAGCTTTGGCACTGACTAGTTCCCTAATAATTCGGTGCC 260
Qy      364 CCGATATGSCCTTCCGGGTATGAAACACCTTCGTTCTGGCTGTTCAATTCGGGTACC 423

```

```

Db      261 CCGATATGCGGTTTCCCGCATAAACAACATAAGCTTCTGACTTCTACCCCTCTCTC 320
Qy      424 GCGATGGCGTGGCTTCGGTGTTCGACCGGGGGTCAAGGTGAGTGGGTTCGGGCGTT 483
Db      321 CTACTCTCTTGCATCTCTATAGTGGAGCGCGGCGAGGAACAGGTTGAACAGTC--- 377
Qy      484 GPTTGGTTCGTGA CCGCGCGCTGTGAC CCGGGAAGCTGGCTATTCGATGGAACCTCGG 543
Db      378 -----TACCCTCCCTTGGCAGGAACTACTCCACCTCGGAGCCTCCGTAGACCTAAAC 431
Qy      544 ATTTTCGCGGTTTCACTTGTCTGGGTGCTCTCGATCATGGGGGCGATCAACATGATCAG 603
Db      432 ATCTTCTCTTACACCTAGAGGTATCTCTCTATCTTAGAGGCAATCAATTTCAACACA 491
Qy      604 ACCTTCTTGAACATGCGCGCCCGCGCATGACCTGCACAAAGTGCCTGTTTCTCGTG 663
Db      492 ACAATTATTAAATAAAACCCCTGCGATAACCAATACCAACAGCCCTTTTCGTCGA 551
Qy      664 TCGATCTTTATCAGGCTTGGCTGATCTCTGGCGCTGCGGTTCTGGCTGCTGCAATC 723
Db      552 TCCGTCTTAATACAGAGCTTACTTCTCTATCTCTCCAGTCTTAGCCGCTGCAATC 611
Qy      724 ACCATGCTGTGACCGACCGTAACCTTCGCGACGACCTTCTTCAATCTCTGTCGGCGGT 783
Db      612 ACTATACTACTACAGACCGTAACCTCAACACCACTTCTTCGACCGCGGAGGAGA 671
Qy      784 GACCGATCTGTACCAACACATCTCTGTTGTTTGGCACCGGGAAGTGTACATCATC 843
Db      672 GACCCCACTTATACCAACACCTTATCTGATTTTTCGGTCACTTCAAGTTTATATCTC 731
Qy      844 ATTCTGCCCGGCTTGGCATCATCAGCCATGTCTGTGCGACCTTCTC-----GAAAAGCC 899
Db      732 ATCTTACGAGCTTCGGAATAATCTCCATATTGTAACCTTACTTACTCCGGGAAAAAAGA 791
Qy      900 GGTCTTCGGTTACTGCGGATGTCTATGCAATGGTGGCAATCGGTGTTCTGGGCTTTGT 959
Db      792 ACCATTTGGATACATAGTATGCTCTGAGCTATGATATCAATTTGGCTTCTCGGGTTAT 851
Qy      960 CGTCTGGGCGCACCATGTACACCGTGTGATGTCTGCTGACCCAGCAATCTCTACTTCAT 1019
Db      852 CGTGTAGCACACCATATATTTACGTAGAAATAGAGTAGACACAGCATATTTCAC 911
Qy      1020 GCTGGCCACCATGTTGATCGCGGTGGCGACCGCATTAAGATCTTCTGTGGATCGCCAC 1079
Db      912 CTCGCTTACCATATATCATCTATCCCAACCGGCTCAAAAGTATTTAGCTGACTCGCCAC 971
Qy      1080 GATGCGGCGGCTCGGTTGAGTTCATATCGCGATCTCTGGGCTTTTGGCTTTATGTT 1139
Db      972 ACTCCACGGAAGCAATGAAATGATCTGCTGAGTCTCTGAGCCCTAGATTTATTTT 1031
Qy      1140 CTTGTTCACTGCTGGTGTGACCGGTATGCTGCTGCCCAAGCGGTTCTGACCGTGC 1199
Db      1032 TCTTTTCACTGCTAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1091
Qy      1200 ATATCAGACACCTATTAAGTGGTGGCGACTTCCATATGATGCTGCTGCTGCTGCTGCT 1259
Db      1092 ACTACAGACACGTAATACGTTGAGCCCACTTCCACTATGCTGCTGCTGCTGCTGCTGCT 1151
Qy      1260 CTTTGGCATCTTCGCGCTATCTACTTTTACATCCGGAAGTTCTCGGCGCGCGCTTTCCC 1319
Db      1152 ATTTGGCATCTATAGAGGCTTCACTGATTTTCCCTATTTCTCAGGCTACACCTAGA 1211
Qy      1320 GGAATGGGCTGCAAAAGCTGCACTTCTGGAAGCTTCTTATCGGTGGGAACGTCAGTTCCT 1379
Db      1212 CCAAACTACGCCAAATCCATTTTCGCTATCATATTCATTCGCGCTAAATCTTAATCTTCT 1271
Qy      1380 CCGGAGCACTTCTGGAAGCTGAGGCTGAGGTTATGCGCGGCTGATGACTATCCCGAAGC 1439
Db      1272 CCCACAAACACTTCTCGGCTATCCGGAATGCCCGAGTGTACTCGGACTATCCCGATGC 1331
Qy      1440 CTTGCGCTGTGGAAACAAAGTCTCTGATGTTGCGGCTCTGCGCTGCTGCTGCTGCT 1499

```


CR615203.1 GI:50496010
HTC; CNSLT CDNA.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1504)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Peng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Faraday Avenue
2 (bases 1 to 1504)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with NotI and cloned
into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
Location/Qualifiers
1. 1504
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODF025YAL9"
/tissue_type="Fetal brain"
/plasmid="pCMVSPORT_6"

Query Match 23.7%; Score 396.2; DB 3; Length 1504;
Best Local Similarity 58.1%; Pred.No.2.9e-95;
Matches 743; Conservative 0; Mismatches 523; Indels 13; Gaps 2;

QY 244 AACGACACCTGTGGAACCTCATGTGTTACTACCATGTTTCTGATGATGTTCTTTTG 303
Db 86 AACGACACATCTACACCTTATCTGTCACAGCCCATGCTTGTAAATATCTTTCATA 145
QY 304 GGTATCCCGCATGTTTGGTGTGTTTGTAACTATCTGATGCGCGTGAATCGCGGT 363
Db 146 GTAATACCATCATATATCGAGGCTTTGGCAACTGACTAGTTCCCTTAATATCGTGCC 205
QY 364 CCGGATATGCGCTTCCCGGTATGAACACCTGTGTTCTGCTGTTTCAATCCCGGTACC 423
Db 206 CCGGATATGCGGTTTCCCGCATAAACACATAGCTTCTGACTCTTACCCCTCTCTC 265
QY 424 GCGATGGGCGTGGCTTTCGCTGTTTCGCACCGGGCGGTGACGGTCAGTGGGTTTCGGCGGT 483
Db 266 CTACTCCTGCTTGCATCTGCTATAGTGGAGGCGCGGCAGAACAGTTGAACAGTC--- 322
QY 484 GGTGGGTTCTGTACCCGCGGTGTCGACCCGGAAGCTGGCTATTCGATGACCTCGG 543
Db 323 -----TACCTCCTTGGCAGGAATACTCCACCTGGAGCTTCCGTAGACCTAAC 376
QY 544 ATTTTCGCGGTTACATGTCGGGTGCTCTCTGATCATGGGCGGATCAACATGATCAG 603
Db 377 ATCTTCTCTTACCTACCTAGCAGGTATCTCTCTATCTTAGGAGCCATCAATTCATCA 436
QY 604 ACTTCTTGAACATGGCGGCCCCCGGATGACGCTGCACAAAGTGCCGTTGTTCGCGG 663
Db 437 ACAATTATTAAATAAAACCCCTGGCATAAACCAATACCAACGCCCCCTTTCGTCTGA 496
QY 664 TCGATCTTTATACCGGTTTGGCTGATCTCTGCTGGCGTCTGGCTGCTGGTGAATC 723
Db 497 TCGTCTCTATACAGCAGTCTTACTTCTCTATCTCTCCAGTCTTAGCCGCTGGCATC 556
QY 724 ACCATGCTGTGACCGACCGTAACTTTCGGCAGCACTTCTTCAATCTCTGCTGGCGCGT 783
Db 557 ACTATACTACTACAGACCGTAACTTCAACACCACTCTTTCGACCCAGCCGAGGAGGA 616

QY 784 GACCCGATTTCTAGCAACACATCTCTGTTCTTTGGGCAACCCGGAAGTGTACATCATC 843
Db 617 GACCCGATTTCTATACCAACACCTATTCTGATTTTCGGTCAACCCCTGAAGTTTATATCTC 676
QY 844 ATTCTGCGCGGCTTTGGCATCATCAGCCATGTCGTGTGACCTTCTC---GAAAGGCC 899
Db 677 ATCTACACGAGCTTCGGAATAATCTCCATATTGTAACTTACTTACTCCGGGAAAAAAGA 736
QY 900 GGTCTTCGGTTACTCGCGATGTCATGCAATGGTGGCAATCGGTGTTCTGGGCTTTGT 959
Db 737 ACCATTGGATACATAGGTATGTCCTGAGCTATGATATCAATGGTCTCTAGGGTTTAT 796
QY 960 CGTCTGGGCGACACATGTACACCGCTGTGATGTGCTGACCCAGCAATCTTACTTTCAT 1019
Db 797 CGTGTGAGCACCATATATTATACGTAGAAATAGACGTAGACACAGGACATATTTCAC 856
QY 1020 GCTGGGCAACATGGTGAATGCGGTGCGGACCGCATTAAGATCTTCTGTTGATCGCCAC 1079
Db 857 CTCGGCTACCATTAATCATCTATATCCCAACCGCGCTCAAGATATTAGCTGACTCGCCAC 916
QY 1080 GATGTGGGCGGCTCGGTTGAGTTCAAAATCGCGATGCTCTGGGCTTTGGCTTTATGTT 1139
Db 917 ACTCCAGGAACAAATAGAAATGATCTGCTGAGTGTCTGAGCCCTAGGATTTATTT 976
QY 1140 CTTGTTTCACTGCGGTGCTGACCGGTATGCTGTCGGCCCAAGCGGTGTGACCGGTGC 1199
Db 977 TCTTTTCACTGAGTGGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1036
QY 1200 ATATACGACACCTATTAATGCTGTCGCGGCTTCCATATGATGTCGCTGGTGGTGGAT 1259
Db 1037 ACTACACGACGCTACTACCTGTTAGCCCACTTCCACTTCTGCTTCAATAGAGGCTGT 1096
QY 1260 CTTTGGCATCTTCCGCGGTATCTATTTTACATGCGCAAGTTCTCGGGCGCGCTTTTCCC 1319
Db 1097 ATTTGCCATCATAGGAGCTTCTATCTGATTTCCCTATTTCTAGGCTACACCTAGA 1156
QY 1320 GGAATGGGCTGAAAGTGCATCTTCTGACCTTCTTCTATCGGTGCGAACCTCAGTTCTT 1379
Db 1157 CCAACCTCATGCGCAAAATCCATTTGCTATCATATTCATCGGGTAAATCTAACTTTCTT 1216
QY 1380 CCGCAGCAGCTTCTCGGAGCTCAGGGTATCGCGGCGGTTTACATGACATATCCCGAAGC 1439
Db 1217 CCCACACATTTCTCGGCTATCCGGAATGCCCGACGTTACTCGGACTATCCGATGC 1276
QY 1440 CTTCCGCTGTGGAACAAAGTCTGCTCTATGTTGGTCTCTGCGCTTCCGCTGCTGCTCT 1499
Db 1277 ATACACACATGAATATCTCTATCATCTGAGGCTCATTTCTTCTAAACAGCAGTAAT 1336
QY 1500 GTTCTTTCATCGTGTCTTT 1518
Db 1337 ATTAATAATTTTCTATAATT 1355

RESULT 15
CR623733
LOCUS full-length cDNA clone CSOCAP002Y018 of Thymus of Homo sapiens 1510 bp mRNA linear HTC 21-JUL-2004
DEFINITION (human).
ACCESSION CR623733
VERSION CR623733.1 GI:50504540
KEYWORDS HTC; CNSLT CDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1510)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Peng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Faraday Avenue

REFERENCE 2 (bases 1 to 1510)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
 COMMENT - Web : www.genoscope.cns.fr)
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
 FEATURES
 source Location/Qualifiers
 1..1510
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0CAP002Y018"
 /tissue types="Thymus"
 /plasmid="pCMVSPORT_6"
 ORIGIN
 Query Match 23.7%; Score 396.2; DB 3; Length 1510;
 Best Local Similarity 58.1%; Pred. No. 2.9e-95;
 Matches 743; Conservative 0; Mismatches 523; Indels 13; Gaps 2;
 QY 244 AACGGACACCTGTGGAACGTCATGGTTACCTACCATGGTATCTCTGATGATGTTCTTTGTG 303
 DB |||||
 QY 77 AACGACACATCTACAACGTTATCGTCAGCCCATGCAATTTGTAATATCTTCTTCATA 136
 DB |||||
 QY 304 GGATATCCCGCATTTGTCGGTGTGTTTGGTAACATATCTGATGCCGTGCAAAATCGCGCT 363
 DB |||||
 QY 137 GTAATACCCATCAATACCGAGGCTTTGGCACTGACTAGTTCCTCCCTAATATATCGGTGC 196
 DB |||||
 QY 364 CCGGATATGCTTCCCGCTATGACAACTCTGCTTGGGTGTCTATGCGGTAC 423
 DB |||||
 QY 197 CCGGATATGCTTCCCGCTATGACAACTCTGCTTGGGTGTCTATGCGGTAC 256
 DB |||||
 QY 424 CGGATGCGGTGCTTGGCTGTTTGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 483
 DB |||||
 QY 257 CTACTCTGCT 313
 DB |||||
 QY 484 GGTGCGGTGCT 543
 DB |||||
 QY 314 -----TACCCCTGCTTGGCAGGAACTACTCCACCCCTGGAGCCTCCGTAGACCTAAC 367
 DB |||||
 QY 544 ATTTCGCGGTTCACCTGTGCT 603
 DB |||||
 QY 368 ATCTTCTCTTACACCTAGCAGTATCTCTCTATCTTAGGAGCCATCAATTTTATCACA 427
 DB |||||
 QY 604 ACCTTCTTGAACTGCGGCGCGCGGATGACGCTGCAAAAGTGCCTGTTGTTCTGCTGG 663
 DB |||||
 QY 428 ACAATTATTAATAAAACCCCTGTCATTAACCAATACCAACGCGCCCTTTTCTGCTGA 487
 DB |||||
 QY 664 TCGATCTTATACCGGTTCGGTGTCTGCT 723
 DB |||||
 QY 488 TCGGTCTTATACAGGAGTCTTACTTCTCTATCTCTCCAGTCTAGCCGCTGCAATC 547
 DB |||||
 QY 724 ACCATGCTGCTGACCGACGTAATCTTGGGACGACCTTCTTCAATCTCTGCGCGCGGT 783
 DB |||||
 QY 548 ACTATACTACTAACAGACCGTAACCTCAACACCACTTCTTCCAGTCTAGCCGAGGAGGA 607
 DB |||||
 QY 784 GACCCGATCTGACCAACACATCTGTTGTTTGGGACCCCGGAGGTGTACATCATC 843
 DB |||||
 QY 608 GACCCGATCTTACCAACACCTTCTGATTTTTCGGTCACTGAGTGTATATTTCTC 667
 DB |||||
 QY 844 ATTCTGCGCGCTTTGGGATCATCAGCCATGTCGTGTCGACCTTCTC-----GAAAAAGCC 899
 DB |||||
 QY 668 ATCTTACCGAGCTTCGGAATATCTCCCATATTTGTAATTTACTACTCCGGGAAAAAAGA 727
 DB |||||
 QY 900 GGTCTTCGTTACTGCGGATGCTCTATGCAATGGTGGCAATCGGTGTTCTCGGCTTTGT 959
 DB |||||
 QY 728 ACCATTGGGATACATAGGTATGCTGCTGAGCTATGATCAATTGGCTTCTTAGGGTTAT 787
 DB |||||

Search completed: October 22, 2004, 18:59:14
 Job time : 7641 secs

QY 960 CGTCTGGCGCACCAATGTATACCGTTGGTATGTGTGCTGACCCAGCAATCCTACTTCAAT 1019
 DB |||||
 QY 788 CGTGTGAGCACACCATATATTTACAGTAGGAATAGACGTAGACACAGGACATATTTTAC 847
 DB |||||
 QY 1020 GCTGGCCACCATGGTATCGCGGTGCCGACCGGATTAAGATCTTCTCGTGATCGGCAC 1079
 DB |||||
 QY 848 CTCGCTACCAATATCATCGGTATCCCGACCGGCTCAAGTATTTAGCTAGCTCGGCAC 907
 DB |||||
 QY 1080 GATGTGGGCGGCTCGGTGAGTTCAATCGCCGATGCTCTGGGCTTTGGCTTTATGTT 1139
 DB |||||
 QY 908 ACTCCACGGAAGCAATATGAAATGATCTGTGAGTGTCTGAGCCCTAGGATTTATTT 967
 DB |||||
 QY 1140 CTTGTTTCACTGGGTGGTGTGACCGGTATGCTGTGCGCCCAAGCGGTCTGGACCGTGC 1199
 DB |||||
 QY 968 TCTTTTCACTGGGTGGTGTGACCGGTATGCTGTGCGCCCAAGCGGTCTGGACCGTGC 1027
 DB |||||
 QY 1200 ATATCAGCACACCTATTACGTGGTGGGCGCACTTCCATTTATGATGCTGCTGGGTGGAT 1259
 DB |||||
 QY 1028 ACTACAGCACAGTACTAGTGGTGGGCGCACTTCCACTATGCTCTATCAATAGGAGCTGT 1087
 DB |||||
 QY 1260 CTTTGGGATCTTGGCGGTATCTACTTTTATCATGCCGAAGTTCTGGGCGCGCTTTCCC 1319
 DB |||||
 QY 1088 ATTTGCCATCATAGGAGGCTTCACTGATTTTCCCTATTTCTCAGGCTACACCCCTAGA 1147
 DB |||||
 QY 1320 GGAATGGGCTGCAAGGCTGCACTTCTGGACCTTCTCATCGGTGCGAAGCTCAAGTTCTT 1379
 DB |||||
 QY 1148 CCAACCTACGCCAATCCATTTGCTATCATATTCTGCGGTAAATCTAACTTTCTT 1207
 DB |||||
 QY 1380 CCGCAGCACTTCTGGGAGCTCAGGATGATGCGCGCGGTATCATCGACTATCCCGAAGC 1439
 DB |||||
 QY 1208 CCGCAACACTTCTCGGCTTATCCGGAATGCCCGGACGTTACTCGGACTATCCCGATGC 1267
 DB |||||
 QY 1440 CTTGCGGTGGAACAAAGTCTGCTCTATGTTGGTCTCTGCGCTTCTGCGCTTCTGCT 1499
 DB |||||
 QY 1268 ATACCAACATGAATATCCATCATCTGAGGCTCATTTCTTCTAAACAGGTAAT 1327
 DB |||||
 QY 1500 GTTCTTTCATGATCTTT 1518
 DB |||||
 QY 1328 ATTAATAATTTTCATAAT 1346
 DB |||||

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 18, 2004, 23:27:21 ; Search time 157 Seconds
(without alignments)
1272.689 Million cell updates/sec

Title: US-09-712-768a-2

Perfect score: 3000

Sequence: 1 MADAAIHGDHHEKQGFTR.....PAHTFTLPKRSWDKHPSH 557

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq 23Sep04:*

1: Genesep1980s:*

2: Genesep1990s:*

3: Genesep2000s:*

4: Genesep2001s:*

5: Genesep2002s:*

6: Genesep2003as:*

7: Genesep2003bs:*

8: Genesep2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3000	100.0	557	4	AA97750 G. oxydan
2	1543	51.4	514	8	ADJ79757 Subunit I
3	1542	51.4	514	7	Ades1901 Rat Prote
4	1531	51.0	513	6	Abg64214 Angiogene
5	1531	51.0	513	8	Adj35456 Human cyt
6	1528	50.9	513	7	Adj68385 Human hea
7	1469.5	49.0	530	6	Abj18768 Pseudomon
8	1278	42.6	425	7	Aboc68563 Pseudomon
9	1223	40.8	573	6	Abu336908 Protein e
10	1223	40.8	573	6	Abu34934 Protein e
11	1215	40.5	574	6	Abu36003 Protein e
12	1192.5	39.8	576	6	Abm66171 Propionib
13	1191.5	39.7	576	6	Abm45147 Propionib
14	1191.5	39.7	576	6	Abm41666 Propionib
15	1187	39.6	543	6	Abu33806 Protein e
16	1175	39.2	564	6	Abu26242 Protein e
17	1169.5	39.0	581	4	Aag92515 C glutam
18	1169.5	39.0	584	5	Aau97041 Coryneb
19	1169.5	39.0	584	5	Abb79439 Coryneb
20	1169.5	39.0	584	7	Add13403 C. Glutam
21	1169.5	39.0	584	8	Ado70036 C glutam
22	1160	38.7	552	4	Aab79569 Coryneb
23	1138.5	38.0	659	5	Abu49819 Listeria
24	1138.5	38.0	659	6	Abu33046 Protein e
25	1098.5	36.6	644	6	Abu18314 Protein e

26	1078.5	35.9	648	4	AAG81983
27	1078.5	35.9	662	6	ABU42645
28	1078.5	35.9	668	5	ABP38585
29	1060.5	35.4	662	6	ABU16562
30	1060.5	35.4	662	6	ABM73311
31	1059.5	35.3	649	4	AAU33837
32	1059.5	35.3	662	4	AAU36750
33	1056.5	35.2	642	6	ABU23408
34	1046.5	34.9	648	6	ABU43487
35	1044	34.8	642	6	ABU16881
36	1044	34.8	664	6	ADA35769
37	1040	34.7	660	6	ABU41038
38	1040	34.7	672	7	ADF05691
39	1023.5	34.1	646	6	ABU22511
40	1020.5	34.0	664	6	ABM67056
41	1019.5	34.0	663	6	ABU50228
42	1010.5	33.7	663	4	AAU34477
43	1010.5	33.7	663	4	AAG98862
44	1010.5	33.7	663	6	ABU31973
45	1010.5	33.7	663	6	ABU28528

ALIGNMENTS

RESULT 1

AA97750

ID AA97750 standard; protein; 557 AA.

XX AC AA97750;

XX DT 06-AUG-2001 (first entry)

XX DE G. oxydans cytochrome C oxidase (COI) protein sequence.

XX KW Cytochrome C oxidase; COI; COII; COIII; COI complex; D-sorbitol; oxidative fermentation; electron transfer; respiratory chain; L-sorbose; 2-Keto-L-gluconic acid; 2KGA production; aldehyde production; carboxylic acid production; ketone production.

XX OS Gluconobacter oxydans.

XX PN EP1103603-A2.

XX PD 30-MAY-2001.

XX PF 14-NOV-2000; 2000EP-00124785.

XX PR 17-NOV-1999; 99EP-00122842.

XX PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX PI Asakura A, Hoshino T, Shinjoh M;

XX XX WPI; 2001-357953/38.

XX DR N-PSDB; AAA91489.

XX PT New cytochrome c oxidase complex having cytochrome c oxidase activity from Gluconobacter oxydans DSM 4025, useful in mediating electron transfer in respiratory chain or producing 2-keto-L-gulonic acid from L-sorbose or D-sorbitol.

XX PS Claim 5; Page 16-20; 42pp; English.

XX CC This sequence is the Gluconobacter oxydans cytochrome C oxidase (COI) of the invention. The COI complex is useful in improving oxidative fermentation and is an essential component mediating electron transfer in the respiratory chain. The recombinant microorganism and the cytochrome C oxidase may be used in the genetic preparation of a recombinant COI complex and in the production of 2-Keto-L-gluconic acid (2KGA) from L-sorbose or D-sorbitol in a culture medium. The COI is also useful as a terminal oxidase, oxidising cytochrome C, an electron acceptor from an enzyme belonging to dehydrogenase for the production of aldehydes,

CC carboxylic acids and ketones from alcohols and aldehydes, especially the
 CC production of 2KGA from L-sorbose or D-sorbitol
 XX
 SQ Sequence 557 AA;

Query Match 100.0%; Score 3000; DB 4; Length 557;
 Best Local Similarity 100.0%; Pred. No. 7.7e-292;
 Matches 557; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MADAAIHGHHHEKQGFTHFSTNKHDIQLLYLVAAGVGVISVLFVYMELELMDPG 60
 DB 1 MADAAIHGHHHEKQGFTHFSTNKHDIQLLYLVAAGVGVISVLFVYMELELMDPG 60

QY 61 VOYMCLEGARLIADASOTCTANGHLNNVMTYHGILMMFFVGIIPALFGGFGNYLPLQIG 120
 DB 61 VOYMCLEGARLIADASOTCTANGHLNNVMTYHGILMMFFVGIIPALFGGFGNYLPLQIG 120

QY 121 APDMAPFRMNNLSFWLPIAGTAMGVSALFAPGQDGLGSGVWLYPPLSTREAGYSMDL 180
 DB 121 APDMAPFRMNNLSFWLPIAGTAMGVSALFAPGQDGLGSGVWLYPPLSTREAGYSMDL 180

QY 181 AIPAVHLSGASSIMGAINMITTLNMEAPGWLHKVPLFSWSIFITAWLILALPLVAGA 240
 DB 181 AIPAVHLSGASSIMGAINMITTLNMEAPGWLHKVPLFSWSIFITAWLILALPLVAGA 240

QY 241 ITMLLTDRNFTGTFPFPAGGDPILYQHILWFFGHPEVYIIILPGFGIISHVVSTFSKPP 300
 DB 241 ITMLLTDRNFTGTFPFPAGGDPILYQHILWFFGHPEVYIIILPGFGIISHVVSTFSKPP 300

QY 301 VFGYLPVYAMVAIGLVGVVAHMYTVGMSLTQOSYFMLATMVAVTGKIFSWIAT 360
 DB 301 VFGYLPVYAMVAIGLVGVVAHMYTVGMSLTQOSYFMLATMVAVTGKIFSWIAT 360

QY 361 MWGGSVEFKSPMLWAFGMFLFTVGGVTGIVLAQAGLDRAYHDTYVVAHFYVMSLGA 420
 DB 361 MWGGSVEFKSPMLWAFGMFLFTVGGVTGIVLAQAGLDRAYHDTYVVAHFYVMSLGA 420

QY 421 FAIPAGIYFYMPKFSGRAPFENAAKLHFMTFFIGANVTFFPQHFLGRQGMRRYIDYPA 480
 DB 421 FAIPAGIYFYMPKFSGRAPFENAAKLHFMTFFIGANVTFFPQHFLGRQGMRRYIDYPA 480

QY 481 FALNNKVSXGAFALPASFLFFVIVFYTLVAGRRETRNPNWGEFADTLEWTLPSPPPAH 540
 DB 481 FALNNKVSXGAFALPASFLFFVIVFYTLVAGRRETRNPNWGEFADTLEWTLPSPPPAH 540

QY 541 TFETLPKRSWDKHPSH 557
 DB 541 TFETLPKRSWDKHPSH 557

RESULT 2

ADJ79757
 ID ADJ79757 standard; protein; 514 AA.

XX AC
 XX AC
 XX ADJ79757;
 DT 06-MAY-2004 (first entry)
 XX
 XX Subunit I of bovine cytochrome c oxidase.
 DE cytochrome c oxidase; slimming drug.
 XX Bos taurus.

XX OS
 XX W02004015095-A1.
 XX PD 19-FEB-2004.
 XX PF 02-JUN-2003; 2003WO-JP006912.
 XX PR 08-AUG-2002; 2002JP-00231042.
 XX PA (NISC-) JAPAN SCI & TECHNOLOGY AGENCY.

XX Yoshikawa S, Shimada H, Shimokata K;
 PI
 XX WPI; 2004-191743/18.
 DR N-PSDB; ADJ79758.

XX Vertebrate cytochrome c oxidase with a modified subunit 1 for potential
 PT use as a slimming drug.
 XX
 XX Disclosure; SEQ ID NO 1; 29pp; Japanese.

XX The present invention relates to a cytochrome c oxidase with a modified
 PS subunit 1 comprising a vertebrate cytochrome c oxidase whose Asp residue
 CC at position 51 from the N-terminal end has been modified. For use in
 CC researching the effects of reduced energy production efficiency in cells
 CC such as nerve and muscle. Potential use as a slimming drug. The present
 CC sequence represents subunit I of bovine cytochrome c oxidase.

XX Sequence 514 AA;

Query Match 51.4%; Score 1543; DB 8; Length 514;
 Best Local Similarity 53.2%; Pred. No. 1.1e-145;

Matches 285; Conservative 87; Mismatches 124; Indels 40; Gaps 7;

QY 17 FFRWFMTNKHDKIGLLYLVAAAGVGVISVLFVYMELELMDPGVQVYMCLEGARLIADAS 76
 DB 2 FINRWLFSTNKHDKIGLLYLVAAAGVGVISVLFVYMELELMDPGVQVYMCLEGARLIADAS 76

QY 77 QFCTANGHLNNVMTYHGILMMFFVGIIPALFGGFGNYLPLQIGAPDMAPFRMNNLSFWL 136
 DB 46 -TLGDDQIYVNVVTAFAVMIFFVMVPIMGSGNWLFLMTGAPDMAPFRMNNLSFWL 104

QY 137 PIAGTAMGVSALFAPGQDGLGSGVWLYPPL--STREAGYSMDLAIPAVHLSGASSIM 194
 DB 105 LPPSFLLLASSMV----EAGAGTGMVYPPPLAGNLHAGASVDLTIFSLHLAGYSSIL 159

QY 195 GAINMITTLNMEAPGWLHKVPLFSWSIFITAWLILALPLVAGAITMLLTDNRNFTGTF 254
 DB 160 GAINFITTIINMKPPASQYQTPLFVMSVMTAVLLLSLVLAAGITMLLTDNRNLTTF 219

QY 255 FNPAGGDPILYQHILWFFGHPEVYIIILPGFGIISHVSTFS-KKPVFGYLPVYAMVA 313
 DB 220 FDPAGGDPILYQHILWFFGHPEVYIIILPGFGIISHVSTFS-KKPVFGYLPVYAMVA 279

QY 314 IGVLFVVAHMYTVGMSLTQOSYFMLATMVAVTGKIFSWIATMVGGSVEFKSPML 373
 DB 280 IGVLFVVAHMYTVGMSLTQOSYFMLATMVAVTGKIFSWIATMVGGSVEFKSPML 339

QY 374 WAFGMFLFTVGGVTGIVLAQAGLDRAYHDTYVVAHFYVMSLGAIFATFAGIYFYMPK 433
 DB 340 WALGFIFLFTVGGVTGIVLAQAGLDRAYHDTYVVAHFYVMSLGAIFATFAGIYFYMPK 399

QY 434 FSGRAFPENAAKLHFMTFFIGANVTFFPQHFLGRQGMRRYIDYPAFALNNKVSXGAF 493
 DB 400 FSGYTLNDTWAKLHFALFMFVGVNMTFFPQHFLGSGMRRYSDYDAYTMWNTISSNGSF 459

QY 494 LAFASFLFFVIVFYTLVAGRRE-----TRNPNWGEFADTLEWTLPSPPPAH 543
 DB 460 ISLTAVMLMVFI-IWEAFASKREVLTVDLTTTN-----LEWLNCGCPPPYHTFE 506

RESULT 3

ADJ61901
 ID ADJ61901 standard; protein; 514 AA.

XX
 XX ADJ61901;
 XX
 XX 29-JAN-2004 (first entry)
 DT
 XX Rat Protein AAB21298, SEQ ID NO 7830.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX OS Rattus norvegicus.
XX FN WO2003016475-A2.
XX PD 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX 14-AUG-2001; 2001US-0312147P.
XX 01-NOV-2001; 2001US-0346382P.
XX 26-NOV-2001; 2001US-0333347P.
XX (GHEO) GEN HOSPITAL CORP.
XX PA (FARB) BAYER AG.
XX PI Woolf C, D'urso D, Befort K, Costigan M;
XX DR WPI: 2003-268312/26.
XX DR GENBANK; AAB21298.
XX
XX New composition comprising two or more isolated polypeptides, useful for
XX PT preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
XX CC or human polynucleotides or a polynucleotide which represents a fragment,
XX CC derivative or allelic variation of the nucleic acid sequence. Also
XX CC claimed are a vector comprising the novel polynucleotide, a host cell
XX CC comprising the vector, a method for identifying a nucleotide sequence
XX CC which is differentially regulated in an animal subjected to pain and a
XX CC kit to perform the method, an array, a method for identifying an agent
XX CC that increases or decreases the expression of the polynucleotide sequence
XX CC that is differentially expressed in neuronal tissue of a first animal
XX CC subjected to pain, a method for identifying a compound which regulates
XX CC the expression of a polynucleotide sequence which is differentially
XX CC expressed in an animal subjected to pain, a method for identifying a
XX CC compound that regulates the activity of one or more of the
XX CC polynucleotides, a method for producing a pharmaceutical composition, a
XX CC method for identifying a compound or small molecule that regulates the
XX CC activity in an animal of one or more of the polypeptides given in the
XX CC specification, a method for identifying a compound useful in treating
XX CC pain and a pharmaceutical composition comprising the one or more
XX CC polypeptides or their antibodies. The polynucleotide or the compound that
XX CC modulates its activity is useful for preparing a medicament for treating
XX CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
XX CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX CC therapy). The sequence presented is a rat protein (shown in Table 2 of
XX CC the specification) which is differentially expressed during pain. Note:
XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic form directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 514 AA;

Query Match 51.4%; Score 1542; DB 7; Length 514;
Best Local Similarity 54.2%; Pred. No. 1.4e-145;
Matches 287; Conservative 84; Mismatches 131; Indels 28; Gaps 7;

QY 17 PFTFPMSTNKHDKIGLLYLVAVAGVGVFISVLFTVMRLMDPDGVQVYMCLEGARLIADAS 76
DB 2 FVNRWLFTSTNKHDKIGLLYLVAVAGVGVFISVLFTVMRLMDPDGVQVYMCLEGARLIADAS 50
QY 77 QTCANGHLNVMVTVHGLIMWFFVGI PALFGGFGNYLMPLOIGAPDMAPPMNNLSFWL 136
DB 51 -----DQYINVIVTAHAFVWFFVWPMVWVGFGNWLVPMLIGAPDMAPPMNNLSFWL 104
QY 137 FIAGTANGVASVAPCGDQGLSGVGVWLYPPL--STREAGYSMDLAFVHLSGASSIM 194
DB 105 LPPSFLLLASSMV-----EAGAGTGWTVYPLAGNLAHAGVSDLTIFSLHAGVSIL 159
QY 195 GAINMITTFLNWRAPGWTLHKVPLFSWSIFITAWLLILLALPLAGAITMLLTDNRFGTTF 254

Db 160 GAINFITTIINKPPAMTQYOTPLFVNSVLIITAVLSSLLSLPVLAAGITMLLTDNRNLTTF 219
QY 255 FNPAGGDDPILYQHILWFFCHPEVYIIILPGFGIISHVVSTFS-KKPVFGYLPWVYAMVA 313
Db 220 FDPAGGDDPILYQHILWFFCHPEVYIIILPGFGIISHVVSTFS-KKPVFGYLPWVYAMVA 279
QY 314 IGVLFVWAHMYTVGMSLTQOSYFNLATMVJAVTGKIPSWIATMGGSGVEFKSPML 373
Db 280 IGFVFWAHHMFTVGLDVTTRAYFTSATMIIAFTGVKFSWLTATLHGGNKNKSPML 339
QY 374 WAFGEMLFTVGGVTVGLVLAQAGLDRAVHDYVYVVAHFYVMSLGAIFAIFAGIYFYMPK 433
Db 340 WALGFILFTVGGTLGIVLSNSSLDIVLHDYVYVVAHFYVMSLGAIFAIFAGIYFYMPK 399
QY 434 FSGRAFPEWAALKLHFWTFFIGANVTFFPQHFLGRQGNPRYIDYPEAFALMNKVSYGAF 493
Db 400 FSGVTLNDTWAKAHFAIMFVGNMTFFPQHFLGLAGNPRYSYDPDAYTTWNTVSMGSGF 459
QY 494 LAFASILFFIVIFVYTLVAGRRETRNPNWGEFADTLEWTLPSPPPAHTPE 543
Db 460 ISUTAVL-VNIFWIEAFASKREVLSISYS--STNLEWLHGCPPPIHIFE 506

RESULT 4
ABR64214
ID ABR64214 standard; protein; 513 AA.
AC ABR64214;
XX 15-OCT-2003 (first entry)
XX Angiogenesis protein BNO108.
XX Cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
XX antipsoriatic; antiarteriosclerotic; cardiac; vasotropic; angiogenesis;
XX gene therapy; vasculature; cancer; rheumatoid arthritis; psoriasis;
XX diabetic retinopathy; cardiovascular disease; atherosclerosis;
XX ischemic limb disease; coronary artery disease.
XX Homo sapiens.
XX WO2003027285-A1.
XX 03-APR-2003.
XX 19-SEP-2002; 2002WO-AU001282.
XX 27-SEP-2001; 2001AU-00007973.
XX 27-SEP-2001; 2001AU-00007974.
XX 11-OCT-2001; 2001AU-00008210.
XX 29-OCT-2001; 2001AU-00008532.
XX 13-NOV-2001; 2001AU-00008838.
XX 28-AUG-2002; 2002AU-00951032.
XX (BION-) BIONOMICS LTD.
XX Gamble JR, Hahn CN, Vadas MA;
XX WPI: 2003-354655/33.
XX N-PSDB; ACP34489.
XX New angiogenic genes and polypeptides, useful for diagnosing, e.g.
XX PT prognosticating or treating an angiogenesis-related disorder, e.g.
XX PT cancer, rheumatoid arthritis, diabetic retinopathy, psoriasis or
XX PT cardiovascular diseases.
XX Claim 15; SEQ ID NO 149; 90pp; English.
XX The invention relates to the isolation of novel genes (ACP34446-ACP34559)
XX CC encoding proteins (ABR64180-ABR64281) involved in the process of
XX CC angiogenesis. The nucleic acid molecules are useful in identifying and/or
XX CC obtaining full-length human genes involved in an angiogenic process. The

nucleic acid molecule, polypeptides or complexes encoded, cells or genetically modified non-human animals derived from these are useful for the screening of candidate pharmaceutical compounds used in treating angiogenesis-related disorders. They are also useful for diagnosing, prognosticating or treating an angiogenesis-related disorder, which involves uncontrolled or enhanced angiogenesis or is a disorder in which a decreased vasculature is of benefit (e.g. cancer, rheumatoid arthritis, diabetic retinopathy, psoriasis or cardiovascular diseases such as atherosclerosis), or involves inappropriately arrested or decreased angiogenesis or is a disorder in which an expanding vasculature is of benefit (e.g. ischemic limb disease or coronary artery disease). The modulator of expression or activity of the polypeptide encoded by the nucleic acid sequence is useful for manufacturing a medicament for the treatment of an angiogenesis-related disorder. This sequence corresponds to one of the novel angiogenic protein

KW	angiogenesis; cytostatic; cancer; gene therapy; human;
KX	cytochrome C oxidase subunit 1; COXI; enzyme.
XX	Homo sapiens.
XX	WO2004019893-A2.
PX	11-MAR-2004.
PD	
XX	
XX	02-SEP-2003; 2003WO-US07523.
XX	
PR	30-AUG-2002; 2002US-00231956.
XX	(RIGE-) RIGEL PHARM INC.
PA	
XX	Lorens JB, Bogenberger J, Holland S, Xu W;
PI	WPI; 2004-239116/22.
XX	N-PSDB; ADL35455.
DR	
DR	
XX	Identifying a compound that modulates angiogenesis for treating e.g.,
PT	cancer comprises contacting the compound with a nucleic acid or
PT	polypeptide and determining the functional effect of the compound upon
PT	the nucleic acid or polypeptide.
XX	
PS	Claim 19; SEQ ID NO 64; 245pp; English.
XX	
CC	The invention relates to a novel method for identifying a compound that
CC	modulates angiogenesis which comprises contacting the compound with a
CC	nucleic acid, or a polypeptide encoded by a nucleic acid, that hybridises
CC	under stringent conditions to a second nucleic acid and determining the
CC	functional effect of the compound upon the nucleic acid or polypeptide.
CC	The method of the invention has cytostatic applications and may be useful
CC	for identifying a compound that modulates angiogenesis for treating
CC	cancer or for use during gene therapy procedures. The current sequence is
CC	that of an angiogenesis-related human protein of the invention.
XX	
SQ	Sequence 513 AA;
Query Match	51.0%; Score 1531; DB 8; Length 513;
Best Local Similarity	52.4%; Pred No. 1.8e-144;
Matches 284; Conservative 81; Mismatches 125; Indels 52; Gaps 6	
QY	17 FFFTRMFSTNHKDIGLLYLVAAGVVGFI SVLFVTVMRLMLDPGVQVWCLEGARLIADAS 76
Db	2 PADRWLFSTNHKDIGTLYLLFGAWAGVLGTALLIRAEILGQPG----- 45
QY	77 QTCATANGHLNMMVTVYHGILLMPEFVGIPALFCGGGHVLMPLQIAGPDMAPFRMNLSFWL 136
Db	46 -NLIGNDHIYNIVITAFAYMIFFVMVPINIMGGFGNALVPLMIGADPMAPPNNKNMSFWL 104
QY	137 FIAGTAGVASLFAFGDGGLSGVGWLVLPPLSTR--EAGYSMDLIAFAVHLSGASSIM 194
Db	105 LPPSLILLLASAMV----EAGAGTGWTVPYPPLAGNVSHPGASVDLTIFSLHLAGVSII 159
QY	195 GAINKITTFLNMRAPMTLHKVLPFSWSFITAILLILPULAGAITMLLTRPFGTTF 254
Db	160 GAINFITTIINMKPPANTQTCTFLVMSVLITAVLSSLPLVLAAGITMLLTRNLNTTF 219
QY	255 FNPAGGGDPLYQHILMFFGHGEPEVI IILPGFGIISHVVSTFS-KKPVFGVLPWVYAMA 313
Db	220 FDPAGGGDPLYQHILMFFGHGEPEVI IILPGFGIISHVVSTFS-KKPVFGVLPWVYAMA 313
QY	314 IGVLGFVWAHHMYTVGMSITQQOSYFMLATMVIAVETGKI KIPSWIATMGGSVEFKSPML 373
Db	280 IGFLGFIWAHHMYTVGMSITQQOSYFMLATMVIAVETGKI KIPSWIATMGGSVEFKSPML 373
QY	374 WAFGFMFLFTVGGVTGLVLAOAGLDRAHDHTYVVAHFHVWVMSLGAI PAIPA GIVFYMPK 433
Db	340 WALGFI FLFTVGGUTGIVLANSSLDIVLHDTYVVAHFHVWVMSLGAI PAIPA GIVFYMPK 433
QY	434 FSGRAFPEAAKHFWTTFFTCANVTFFPQHFLGRQCQMRRYIDYDEAFALWNKKVSVSGAF 493

Db 400 FSGYTLDTYAKIHFTIMFIGVNLTPFPQHLGSGMPRRYSYDPDAYTTWNILSSVGSF 459
 QY 494 LAFASFLFFIVFYVTLVAGRRTRPNPWGEFAD-----TLEWTLPSPPAHT 541
 Db 460 ISLRAVLMIMFI-----WEAFASKRKVLMEESMNLWLYGCPPEYHT 504
 QY 542 FE 543
 Db 505 FE 506
 RESULT 6
 ADJ68385
 ID ADJ68385 standard; protein; 513 AA.
 AC ADJ68385;
 XX
 XX 06-MAY-2004 (first entry)
 XX
 XX Human heat mitochondrial protein as a therapeutic target SeqID191.
 XX
 XX mitochondrial; human; screening assay; diabetes mellitus;
 KW Huntington's disease; osteoarthritis;
 KW Leber's hereditary optic neuropathy; LHON;
 KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
 KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
 KW neuroprotective; nontropic; antidiabetic; anticonvulsant; antiarthritic;
 KW osteopathic; ophthalmological; cytostatic.
 XX
 XX Homo sapiens.
 OS
 XX WO2003087768-A2.
 XX
 XX 23-OCT-2003.
 XX
 XX 04-APR-2003; 2003WO-US010870.
 XX
 XX 12-APR-2002; 2002US-0372843P.
 XX
 XX 17-JUN-2002; 2002US-036987P.
 XX
 XX 20-SEP-2002; 2002US-0412418P.
 XX
 XX (MITO-) MITOKOR.
 PA
 PA (BUCK-) BUCK INST AGE RES.
 XX
 XX Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
 PI Warnock DE;
 XX
 XX WPI; 2003-845369/78.
 DR
 XX Identifying a mitochondrial target for drug screening assays and for
 PT treating diseases associated with altered mitochondrial function,
 PT comprises detecting a modified polypeptide in a sample and correlating
 PT with the disease.
 XX
 XX Claim 1; SEQ ID NO 191; 180pp; English.
 PS
 XX This invention relates to novel mitochondrial targets that can be used
 CC for therapeutic intervention in treating a disease associated with
 CC altered mitochondrial function. Specifically, it refers to a method for
 CC identifying proteins of the human heart mitochondrial proteome that are
 CC useful for drug screening assays, as well as therapeutic targets. The
 CC present invention describes a method for identifying such proteins that
 CC can be used in the treatment of various diseases associated with altered
 CC mitochondrial function including diabetes mellitus, Huntington's disease,
 CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
 CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
 CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
 CC compositions have neuroprotective, nontropic, antidiabetic,
 CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
 CC cytostatic activities. This polypeptide sequence is a human heart
 CC mitochondrial protein of the invention.
 XX
 XX Sequence 513 AA;
 SQ

Query Match 50.9%; Score 1528; DB 7; Length 513;
 Best Local Similarity 52.4%; Pred. No. 3.7e-144;
 Matches 284; Conservative 81; Mismatches 125; Indels 52; Gaps 6;
 QY 17 FFRWFNSNHNKDIGLLYLVAAGVGVFISVLFTVYMRLEMDPGVQVCMLEGARLIADAS 76
 Db 2 FADRWLFSTNHNKDIGTLYLLFGNAWGLGTALSLRLRABLGQPG-----45
 QY 77 QICTANGHLNVMVTVYHGLIMMFFVVGIPALFGSGNYLMLPQIGAPDAPPRNNILSFWL 136
 Db 46 -NLGNDHIYVIVTAHAFWMIFFWMPIMIGFGNWLVPVLMIGAPDAPPRNNILSFWL 104
 QY 137 FIAGTANGVASLAPGDDGOLSGVGVWLYPPPLSTR--EAGYSMDLAIFAVHLSGASSIM 194
 Db 105 LPESLILLLLASAMV-----EAGAGTGTWYPPPLAGNYSHPGASVDLTI FSLHLAGVSSIL 159
 QY 195 GAINMITTFLNMRAPGMTLHKVPLFWSIFITAWLILLALPVLAGAITMLLTDNFGTTF 254
 Db 160 GAINFITTIINMKPPANTQYQPLFWNSVLIITAVLLLSLPVLAAGITMLLTDNLTTF 219
 QY 255 FNPAGGDPILYOHILWFFGHPEVYIILPGGIIISHVVSTFS-KXPVFGYLPWYAMVA 313
 Db 220 FDPAGGDPILYQHLSWFFGHPEVYIILPGFGMISHIVTYYSKKEPFYMGWMMMS 279
 QY 314 IGVLFVWVAHMYTVGMSITQOSYFMLATMVAIVPTGKIPSWIATMWSGVSVERKSPML 373
 Db 280 IGFGLFVVAHMYTVGMDVDTTRAITTSMTIITAIPIGVKVSFWLALHGSNNKNSAAVL 339
 QY 374 WAFGFMFLFTVGGVTGIVLAQAGLDRAVDYTVVAHMYVMSLGAIFAIFAGIYFMPK 433
 Db 340 WALGFIFLFTVGGTGLTIVLANSLSLDVLDHYTVVAHMYVMSLGAIFAIFAGIYFMPK 399
 QY 434 FSGRAPEWNAKLHFWTFPGANVTFFPQHFLGQCMPEYIDYDEAFALNKKVSVGAF 493
 Db 400 FSGYTLDTYAKIHFAIMFIGVNLTPFPQHLGSGMPRRYSYDPDAYTTWNILSSVGSF 459
 QY 494 LAFASFLFFIVFYVTLVAGRRTRPNPWGEFAD-----TLEWTLPSPPAHT 541
 Db 460 ISLTAVMLMIFMI-----WEAFASKRKVLMEESMNLWLYGCPPEYHT 504
 QY 542 FE 543
 Db 505 FE 506
 RESULT 7
 ABJ18768
 ID ABJ18768 standard; protein; 530 AA.
 AC ABJ18768;
 XX
 XX 27-FEB-2003 (first entry)
 DT
 XX Pseudomonas aeruginosa biofilm formation-related protein #32.
 DE
 XX Biofilm formation modulation; biofilm-associated disease;
 KW cystic fibrosis; AIDS; middle ear infection; acne; periodontal disease;
 KW catheter-associated infection; medical device-associated infection.
 XX
 XX Pseudomonas aeruginosa.
 OS
 XX WO2000285295-A2.
 PN
 XX 31-OCT-2002.
 PD
 XX 19-APR-2002; 2002WO-US012532.
 PF
 XX 20-APR-2001; 2001US-0295190P.
 XX
 XX 24-OCT-2001; 2001US-0344142P.
 PR
 XX (IOWA) UNIV IOWA RES FOUND.
 PA
 PA (HARD) HARVARD COLLEGE.

XX Whiteley M, Bangera MG, Lory S, Greenberg EP;
 XX WPI; 2003-075601/07.
 DR N-PSDB; ABT14590.
 XX Identifying compound capable of modulating biofilm formation by
 PT bacteria/bacterial antibiotic resistance, useful for treatment of biofilm
 PT associated disease.
 XX Claim 1; Page 116-118; 154pp; English.
 XX The invention comprises a method for identifying a compound capable of
 CC modulating biofilm formation by bacteria. The method of the invention is
 CC useful for identifying a compound capable of modulating biofilm formation
 CC by bacteria or modulating bacterial antibiotic resistance. The method of
 CC the invention is also useful for diagnosing and treating a subject
 CC (especially an immunocompromised human) that is afflicted with a biofilm-
 CC associated disease or disorder, such as: cystic fibrosis; AIDS; middle
 CC ear infections; acne; periodontal disease; catheter-associated infections
 CC ; and medical device-associated infections. The present amino acid
 CC sequence represents a protein that is used in the invention
 XX
 SQ Sequence 530 AA;

Query Match 49.0%; Score 1469.5; DB 6; Length 530;
 Best Local Similarity 51.4%; Pred. No. 2.8e-138;
 Matches 279; Conservative 91; Mismatches 146; Indels 27; Gaps 8;
 QY 7 HGHDDHE--KQGFTRFMSNTHKDTGLLYLVAGVGVISULFTVYMWLELMDPGVQM 64
 DB 11 HAGDHGHPAKGLM-RVLTNTHKDTGLTYLWFSFMMFLLOGSMAMVIAEALFQFGLQ-- 67
 QY 65 CLECARLIADASQTCANGHLNMVMTYHILMMFVFGIPALFGGNGYLMPLQIGAPDM 124
 DB 68 -----IVEFA-----FFNQMTTHGLINVFCANVFA-FVGLANWMIPLMIGAPDM 111
 QY 125 AFRPMNLSFWLFIAGTANGVASLAPGGDGLGSGVGVWLYPPPLSTREAGYSMDLAIPA 184
 DB 112 ALPRMNFSLPAAFGLLVSTLFEPGG---GNFGMTFYAPLSTTFAPHSVTFEIPA 167
 QY 185 VHLGSSAIMGAINMTITPLNNRPAQMTLHKVPLRFSWIFITAWLILLALPVLGAITML 244
 DB 168 IHLGSSIMGAINVATILNLRAPGXTLMKPLFWMTLITAFLLIAMPVLAGVITNM 227
 QY 245 LTRNFGTTFNPPAGGDPILYQHILWFGHPBEVYIIILPGFGIISHVVSTSKPVPFGY 304
 DB 228 LMDIHFGTSFSAAGGDPVLFQFVFWFFGHPBEVYIMILPAFGAVSAIITPARKPLFGY 287
 QY 305 LPMVYAWAIVGLGVFWVAHHMYTVGMSLTQOSYFNLATMTVIAVPTGKIFSWIATMGG 364
 DB 288 TSMVYATASIAFLSFWVAHHMFVVGIPVTGELFFMTATMLIAVPTGVKVFNNVTMMEG 347
 QY 365 SVFEKSPMLWAFGMFLFTVGGVTIGVLAQAGLDRAHYDITYVVAHFHYVMSLGAIFAIF 424
 DB 348 SLTFETPMLFAVAVILFTIGGSGMLAIPADFOYHDTYFVAHFHYVLPVGAIFGIF 407
 QY 425 AGIYFYKPKSGRAPEWAKLHFWTFIFGANTVFPFQHLGROGMPRIYIDYPAFALW 484
 DB 408 ASAYYLPKWTGHYDETILGKLHFWMSFTGMNLAFEPFMHFVGLAGMPRIPIPNYQFADF 467
 QY 485 NKVSSYGAFIAPASLFFFIYFVTVLVAQRRTNPNMGEFADTLEWLPSPPPAHTPET 544
 DB 468 NMVSSIGAFNFGTQLLFLFI-VKIRGGKAPAKPW-DGAGLEWSIPSPAPYHTFST 525
 QY 545 LPK 547
 DB 526 PPE 528
 RESULT 8
 ABO68563
 ID ABO68563 standard; protein; 425 AA.

XX ABO68563;
 XX 29-JUL-2004 (first entry)
 XX Pseudomonas aeruginosa polypeptide #738.
 XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
 XX Pseudomonas aeruginosa.
 XX US6551795-B1.
 XX 22-APR-2003.
 XX 18-FEB-1999; 99US-00252991.
 XX 18-FEB-1998; 98US-0074788P.
 XX 27-JUL-1998; 98US-0094190P.
 XX (GENO-) GENOME THERAPEUTICS CORP.
 XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
 XX WPI; 2003-615309/58.
 XX N-PSDB; ABD02134.
 XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.
 XX Disclosure; SEQ ID NO 17309; 455pp; English.
 XX The invention relates to Pseudomonas aeruginosa polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-P. aeruginosa drugs, as templates for recombinant
 CC production of P. aeruginosa-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of P. aeruginosa-caused
 CC infection, and in detection of P. aeruginosa sequences or other sequences
 CC of Pseudomonas species using biochip technology. Sequences ABO67826-
 CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html
 XX
 SQ Sequence 425 AA;

Query Match 42.6%; Score 1278; DB 7; Length 425;
 Best Local Similarity 55.6%; Pred. No. 3.5e-119;
 Matches 234; Conservative 71; Mismatches 110; Indels 6; Gaps 3;
 QY 127 PRMNLSFWLFIAGTANGVASLAPGGDGLGSGVGVWLYPPPLSTREAGYSMDLAIFAVH 186
 DB 9 PRMNFSFWLLPAAFGLLVSTLFEPGG---GNFGMTFYAPLSTTFAPHSVTFEIPA 64
 QY 187 LSGASSIMGAINMTITPLNNRPAQMTLHKVPLRFSWIFITAWLILLALPVLGAITMLT 246
 DB 65 LAGTSSIMGAINVATILNLRAPGXTLMKPLFWMTLITAFLLIAMPVLAGVITMMLM 124
 QY 247 DRNFGTTFNPPAGGDPILYQHILWFGHPBEVYIIILPGFGIISHVVSTSKPVPFGYLP 306
 DB 125 DIHFGTSFSAAGGDPVLFQFVFWFFGHPBEVYIMILPAFGAVSAIITPARKPLFGYTS 184
 QY 307 MYVATASIAFLSFWVAHHMYTVGMSLTQOSYFNLATMTVIAVPTGKIFSWIATMGGSV 366
 DB 185 MYVATASIAFLSFWVAHHMFVVGIPVTGELFFMTATMLIAVPTGVKVFNNVTMMEGSL 244
 QY 367 EFKSPMLWAFGMFLFTVGGVTIGVLAQAGLDRAHYDITYVVAHFHYVMSLGAIFAIFAG 426

Db	245	TTETPNLFAVAFLVETTTGGFSGMLAIAPADFQYHDTYFWAHFYVLVFGAIFGIPAS	304
Qy	427	IVFYWPKFSGRAFPSPWAAKLHWTFIFFIGANTYFFPQHFGLRGQMPRRYIDYPEAFALMNK	486
Db	305	AYYWLFPKWTHGYDETTLGKLHFWMSFIGMNLAFPFMHFVGLGNPRRI PDYNIQLQFADFNM	364
Qy	487	VSSYGNLAFASFLPIYIVFYVYLLVAGRRRETPNKPGEPAADLTLEWTLPSPPPAHTFETLP	546
Db	365	VSSIGAFMGTTQLLFLFI-VIKIRGGPAPAPKW-DGAEGLSNSIPSPAPHTFTSTTP	422
Qy	547	K	547
Db	423	E	423

RESULT 9
ABU36908
ID ABU36908 standard; protein; 573 AA.
XX
XX AC ABU36908;
XX
DT 19-JUN-2003 (first entry)

CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX
S Sequence 573 AA:

RESULT 10
ABU34934
ID ABU34934 standard: protein: 573 AA.

XX
DT 19-JUN-2003 (first entry)

XX DE protein encoded by Prokaryotic essential gene #20461.

Antisense: prokaryotic essential gene; cell proliferation; drug design.

Claim 25; SEQ ID NO 63927; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *X. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 574 AA;

Query Match 40.5%; Score 1215; DB 6; Length 574;
Best Local Similarity 43.7%; Pred. No. 1.1e-112;
Matches 237; Conservative 94; Mismatches 177; Indels 34; Gaps 9;

QY 24 STNHKDIGLLYLVAAGVGIFSVLFTVMRLBMDPGVQVCMLEGARLIADASQCTTANG 83
DB 34 TTDHROMIGIMYCVACPIFFVFGGLLALLMRTLAAPGLQLF-----SN 76
QY 84 HLWNVMVTYHGLTMMPFVCI PALFGFGNLYMPLQIGAPDMPAPRNMLNSFWLFIAGTAM 143
DB 77 EQFNQLFTWGHGIMLLFYATPIVF -GFANLVLPQLGAPDVAFPRLNAPSWFLFPGA 135
QY 144 GVASLFAFGGDLGQSGVGWVLYPPL--STRAGYSMDLAI FAVHLSGASSIMGAINMIT 201
DB 136 GNAGFITPGG---AADFGWTAYTDLTAHSPGVGGDLWIMGLVAGLGTILGAVNMIT 191
QY 202 TELNMRAPGWTLEKVPLEFSNIFITAWLJLLALPVLGAGITMLLTDNRGTTFFNPAGGG 261
DB 192 TVVCLSAFGMTWFMPIFTWNLIVTSILVLIAPLLTAALFGLADRHGAHIYDAANGG 251
QY 262 DFLYQHILWFFGHPEVYIILPGFGII SHVSVSTESKPEVFGYLPVMYAMVAIGLVGV 321
DB 252 -VLLWQHLEWFFGHPEVYIILPFGGIVSEIPEVRSKKEIFGVTTLVYATLSIAALSVA 310
QY 322 WAHMTVGMSLTQQSFYMLATWVIAPVTGIKI FSWIATWVGSGVEFKSPMLWAGFMFL 381
DB 311 WAHMTFATGAVLL--PFFSMTYLLIAPVTGIKFNNWVGTMWKGQUTFTFPMFLSVGF 368
QY 382 FTVGGVTVGLVLAAGLDRAVDHYVYVAHFVYVMSLGAIFALFAGIYFYMPKFSGRAPPE 441
DB 369 FLGGTLGVLASPPDLDFHVDTSYFVVAHFVYVLFCTIVFSFAGIYFWFPMWTRGLLDE 428
QY 442 WAAXLHFWTFIGANTYFPQHLFGQGMPPRYIDY--PEAFALMNKVSYSYGAFLAFASF 499
DB 429 QLGLKHLFWLTFIGFHTTFLVQHLWGMGMPPRYADYLDPTDGFQGLNWNVTGSGFILGSM 488
QY 500 LRFIVTVVTLVAGRETRPNWGEFADTLEWTLRSPPPAHTFETLPK---RSDWDKHP 555

DB 489 PFPVWNVKSWRGYGVTVDDPMWG-YGNSLEWATSCPPRRHNFTELPRISRPAFELHY 547
QY 556 SH 557
DB 548 PH 549

RESULT 12

ABM66171

ID ABM66171 standard; protein; 576 AA.

XX AC ABM66171;

XX 20-OCT-2003 (first entry)

XX Propionibacterium acnes immunogenic polypeptide #30847.

XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
immunostimulant; immune response; vaccine; immunogenic.

XX Propionibacterium acnes.

XX WO2003033515-A1.

XX 24-APR-2003.

XX 11-OCT-2002; 2002WO-US032727.

XX 15-OCT-2001; 2001US-00978825.

XX (CORI-) CORIXA CORP.

XX Mitcham JL, Steiky YAW, Persing DH, Bhatia A, Maisonneuve JL;

PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;

PI Barth Y, Vallie-Douglass J;

XX WPI; 2003-381789/36.

XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
or for stimulating an immune response specific for a *P. acnes* protein.

PS Claim 7; SEQ ID NO 30847; 1481pp; English.

XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
encoding a Propionibacterium acnes protein. The invention also relates to
polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
immunogenic fragments of *P. acnes* polypeptides. The invention
additionally encompasses expression vectors and host cells comprising a
polynucleotide of the invention; antibodies against polypeptides of the
invention; fusion proteins comprising a polypeptide of the invention; a
method for stimulating an immune response specific for a *P. acnes*
polypeptide and an isolated T cell population comprising T cells prepared
via this method; a vaccine composition (comprising *P. acnes* polypeptides,
polynucleotides, antibodies, fusion proteins, T cell populations, or
antigen-presenting cells that express the polypeptide); a method and kit
for detecting or determining the presence or absence of *P. acnes* in a
patient; and a method for inhibiting the development of *P. acnes* in a
patient. The *P. acnes* polypeptides, polynucleotides, antibodies, fusion
proteins, T cell populations or antigen-presenting cells that express the
polypeptides are useful for diagnosing, preventing or treating acne
vulgaris, or for stimulating an immune response specific for a *P. acnes*
protein. The polynucleotides can also be used as probes or primers for
nucleic acid hybridisation. The vaccine composition is useful for the
stimulation of an immune response against *P. acnes*, or for treating acne,
and the kit is useful for performing a diagnostic assay. The present
sequence represents a specifically claimed *P. acnes* polypeptide which is
thought to contain an immunogenic region. Note: The sequence data for
this patent did not form part of the printed specification, but was
obtained in electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences

PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

PS Claim 25; SEQ ID NO 61730; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 543 AA;

Query Match 39.6%; Score 1187; DB 6; Length 543;
Best Local Similarity 43.8%; Pred. No. 6.6e-110;
Matches 235; Conservative 86; Mismatches 181; Indels 34; Gaps 9;
QY 30 IGLLYVAAGVGVFISLVFTVYMELEMDPGVQVCMLEGARLIADASQTCTANGHLWNVM 89
DB 2 IGIMYVTCRAFFIGGIMALLMRTELAAPGLQL-----SNEQFNOL 44
QY 90 VYHGLIMFEVGIPLPFGGNGVNLPLQIGAPDMPAPRMNLSFWLFIAGTAMGVASLF 149
DB 45 FTWEGTIMLLLYATPVVF-GFANLVLPLOIGAPDVAFFRLNAPFWLFLFGGLIASGFI 103
QY 150 APGGDGLGSGVGWVLPPLS--TREAGYSMDLAFVHLSGASSIMGAINMTTFLNMR 207
DB 104 VPGG----ADFGWTATPLSDAVHSPGAGGDLWITGLVAGLGTILGAVNMITTVCNR 159
QY 208 APGMLHKVPLFNSWIFITAWLILLALPVLAGAITMLLTDRNPGTTFFNPAGGDPILYQ 267
DB 160 APGWTMRMEPIFTWNTILVTSILIIAFTIITLALFGLAADRHLAGHVYDAANGG-VLLWQ 218
QY 268 HILWFFGHPEVYIIILPFGGIIISHVSTFSKPVGCLPMYVAMVAIGLVGVWAHMY 327
DB 219 HLFWFFGHPEVYIIALPFGGIIISHVSTFSKPVGCLPMYVAMVAIGLVGVWAHMY 278
QY 328 TVGMSLTQSQSYFMLATVIAVPTGIKIFSWIATMGGSVFEKSPMLWAFGMFLFTVGCV 387
DB 279 ATGAVLL--PFGSMYLIAPVTGIKFENWIGTMWKGQLTEETMLFCVGFLLTLLGGL 336
QY 388 TGIVLAQAGIDRAYHTYVVAHFYVMSIGATFAIFAGIYFYMPKFSGRAPENAKLH 447
DB 337 TGWLASPPLDHVTDTYFVVAHFYVLPFGTIVFATFAGVYFWPMTGRLLDERIGKLH 396
QY 448 FWTFPGANVTFFPQHFLGSGQPRYIDY--PEAFALWNKVSSYGAFALAFASFLFFVI 505

DB 397 FWLTFIGFHTTFLVOHWLGDLMGPRRYADYLPDSGFOFYNVASTVGAFILGASMFPPVWN 456
QY 506 FVYTLVAGRRRETRNPNFPGCEADTLEWTLPSPPPAHTTETLPK---RSDMDKHPSH 557
DB 457 VPKSHRYGEVTVDDFWG-YGNSLEWATSCPPPHNTTELPRISRPAFELHYPH 511

Search completed: October 18, 2004, 23:41:10
Job time : 162 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 18, 2004, 23:16:05 ; Search time 41 Seconds
(without alignments)
900.954 Million cell updates/sec

Title: US-09-712-768A-2
Perfect score: 3000
Sequence: 1 MADAAIHGDHHEKQGFTR.....PAHTFETLPKRSQWDKHPH 557

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pap:*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pap:*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pap:*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pap:*
5: /cgn2_6/ptodata/1/iaa/PTUS.COMB.pap:*
6: /cgn2_6/ptodata/1/iaa/backfiles.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1531	51.0	513	3	US-09-097-889-15 Sequence 15, Appl
2	1531	51.0	513	4	US-09-098-079-15 Sequence 15, Appl
3	1278	42.6	425	4	US-09-252-991A-17309 Sequence 17309, A
4	1078.5	35.9	648	4	US-09-710-279-1060 Sequence 1060, Ap
5	1078.5	35.9	668	3	US-09-134-001C-3430 Sequence 3430, Ap
6	1044	34.8	664	4	US-09-328-352-7056 Sequence 7056, Ap
7	1040	34.7	672	4	US-09-543-681A-5976 Sequence 5976, Ap
8	1010.5	33.7	663	4	US-09-711-164-332 Sequence 332, App
9	1010.5	33.7	677	4	US-09-489-039A-13088 Sequence 13088, A
10	974.5	32.5	660	4	US-09-252-991A-22396 Sequence 22396, A
11	234.5	7.8	98	4	US-09-248-796A-16350 Sequence 16350, A
12	186	6.2	472	3	US-09-354-123-2 Sequence 2, Appli
13	186	6.2	472	4	US-09-504-357-2 Sequence 2, Appli
14	163	5.4	161	4	US-09-248-796A-15595 Sequence 15595, A
15	155.5	5.2	474	4	US-09-252-991A-30354 Sequence 30354, A
16	151.5	5.1	423	4	US-09-540-236-2311 Sequence 2311, Ap
17	151.5	5.1	766	4	US-09-248-796A-16626 Sequence 16626, A
18	151	5.0	63	4	US-09-248-796A-14194 Sequence 14194, A
19	150	5.0	623	4	US-09-252-991A-19867 Sequence 19867, A
20	149.5	5.0	538	4	US-09-252-991A-23060 Sequence 23060, A
21	148.5	5.0	61	4	US-09-248-796A-15596 Sequence 15596, A
22	145.5	4.9	501	4	US-09-252-991A-28456 Sequence 28456, A
23	130.5	4.3	507	4	US-09-540-236-3391 Sequence 3391, Ap
24	130.5	4.3	521	4	US-09-540-236-2193 Sequence 2193, Ap
25	128.5	4.2	484	4	US-09-107-532A-3830 Sequence 3830, Ap
26	125	4.2	345	4	US-09-252-991A-23059 Sequence 23059, A
27	125	4.2	417	4	US-09-489-039A-8565 Sequence 8565, Ap

ALIGNMENTS

RESULT 1

US-09-097-889-15

Sequence 15, Application US/09097889

Patent No. 6218117

GENERAL INFORMATION:

APPLICANT: Heinstadt, Corrina

APPLICANT: Ghosh, Soumitra S.

APPLICANT: Davis, Robert E.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING

TITLE OF INVENTION: AGENTS THAT QUANTITATIVELY ALTER DETECTABLE

TITLE OF INVENTION: EXTRAMITOCHONDRIAL DNA: MITOCHONDRIAL DNA RATIOS

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED AND BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA: US/09/097,889

APPLICATION NUMBER: 15-JUN-1998

FILING DATE: 15-JUN-1998

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Rosenman Ph.D., Stephen J.

REGISTRATION NUMBER: 43,058

REFERENCE/DOCKET NUMBER: 660088.417

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 682-6031

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 513 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-09-097-889-15

Query Match 51.0%; Score 1531; DB 3; Length 513;
Best Local Similarity 52.4%; Pred No. 2.9e-142;
Matches 284; Conservative 81; Mismatches 125; Indels 52; Gaps 6;

QY 17 FFFRWFMTNKHKIGLLYLAAGVGVGIFSVLFTVYMLELMDPGVQVNCLEGARLIADAS 76
DB 2 FADRWFSTNKHKIDIGTLLYLLFGAWAGVLGTALSLLIRAEALGQPG----- 45

QY 77 QTCTANGHLNVMVYHYHILMMFFVIGIPALFGGNGVLMPLQIGAPDMAPPRNNLSFWL 136
DB 46 -NLLGNDRHIVNIVTAHAFVIMFFVMPIMIGFGNWLVPMLGAPDMAPPRNNNSFWL 104
QY 137 FIAGTAMGVASLFAFGGDLGSGGVGVLYPLSTR--EAGYSMDLAIFAVHLSGASSIM 194
DB 105 LPSLLILLASAMV-----EAGAGTGMVYPPPLAGNYSHPGASVDLTIIFSLHLAGVSSIL 159
QY 195 GAINMHTTFLNMRAPGTLHKVPLFSWISIFITAWLILLALPVLGAIITMLLTDNRFGTTF 254
DB 160 GAINFIITIIINMKPPATQOTPLFVMSVLITAVLLLSLPLVLAAGITMLLTDNRNLNTTF 219
QY 255 FNPAGGDDPILYQHILWFFHGFHPEVYIIILPGFGIISHVWSTFS--KKPVFGYLPVMYAMVA 313
DB 220 FDPAGGDDPILYQHILWFFHGFHPEVYIIILPGFGMISHIVYYSKKPEFGYMGVWAMVS 279
QY 314 IGVLFVWAHMYTVGMSLTQOSYFMLATWIAVPTGKIFSWIATMGGSVFEKSPML 373
DB 280 IGVLFVWAHMYTVGMSLTQOSYFMLATWIAVPTGKIFSWIATMGGSVFEKSPML 373
QY 374 WAFGMFLFTVGGVTGIVLAQAGLDRAVHDTYVVAHFHYVMSLGAIFAIFAGIYFMPK 433
DB 340 WALGFIFLFTVGGTGLIVLANSLSLIDVLDHTYVVAHFHYVLSMGAVFAIMGFIHWFPL 399
QY 434 FSGRAPPEWAAKLHFWTFFIGANVTFFPQHFLGRQGMPPRYIDYPAFALWNVKVSYGAF 493
DB 400 FSGYTLDDQTYAKIHFTIMFIGNLTFFPQHFLGSGMPRRYSYDPAITWNILSSVGSF 459
QY 494 LAFASLFFIIVFYVTLVAGRRTNPNMGEFAD-----TLEWTLPSPPAHT 541
DB 460 ISLTAVMLMIFMI-----WEAFASKKVLWMBEPPSNMLEWLYGCPPPYHT 504
QY 542 FE 543
DB 505 FE 506

RESULT 2

US-09-098-079-15
; Sequence 15, Application US/09098079
; Patent No. 6489095
; GENERAL INFORMATION:
; APPLICANT: HerinStadt, Corrina
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Cleveland, William
; APPLICANT: Fahy, Eoin F.
; APPLICANT: Davis, Robert E.
; TITLE OF INVENTION: DIAGNOSTIC METHOD BASED ON QUANTIFICATION OF
; TITLE OF INVENTION: EXTRAMITOCHONDRIAL DNA
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/098,079
; FILING DATE: 15-JUN-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rosentman Ph.D., Stephen J.
; REGISTRATION NUMBER: 43,058
; REFERENCE/DOCKET NUMBER: 660088.416
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 15:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 513 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-09-098-079-15

Query Match 51.0%; Score 1531; DB 4; Length 513;

Best Local Similarity 52.4%; Pred. No. 2.9e-142;

Matches 284; Conservative 81; Mismatches 125; Indels 52; Gaps 6;

QY 17 FTFWSTHSTHDKDGLLVLAAGVGVGISVLTFTVVMLELMMDPGVQVMCLGARIADAS 76

DB 2 PADRWLSTNHKDI GTLYLLFGAWAGVLTALSLLIRAEIGQPG----- 45

QY 77 QTCTANGHLNVMVYHYHILMMFFVIGIPALFGGNGVLMPLQIGAPDMAPPRNNLSFWL 136

DB 46 -NLLGNDRHIVNIVTAHAFVIMFFVMPIMIGFGNWLVPMLGAPDMAPPRNNNSFWL 104

QY 137 FIAGTAMGVASLFAFGGDLGSGGVGVLYPLSTR--EAGYSMDLAIFAVHLSGASSIM 194

DB 105 LPSLLILLASAMV-----EAGAGTGMVYPPPLAGNYSHPGASVDLTIIFSLHLAGVSSIL 159

QY 195 GAINMHTTFLNMRAPGTLHKVPLFSWISIFITAWLILLALPVLGAIITMLLTDNRFGTTF 254

DB 160 GAINFIITIIINMKPPATQOTPLFVMSVLITAVLLLSLPLVLAAGITMLLTDNRNLNTTF 219

QY 255 FNPAGGDDPILYQHILWFFHGFHPEVYIIILPGFGIISHVWSTFS--KKPVFGYLPVMYAMVA 313

DB 220 FDPAGGDDPILYQHILWFFHGFHPEVYIIILPGFGMISHIVYYSKKPEFGYMGVWAMVS 279

QY 314 IGVLFVWAHMYTVGMSLTQOSYFMLATWIAVPTGKIFSWIATMGGSVFEKSPML 373

DB 280 IGVLFVWAHMYTVGMSLTQOSYFMLATWIAVPTGKIFSWIATMGGSVFEKSPML 373

QY 374 WAFGMFLFTVGGVTGIVLAQAGLDRAVHDTYVVAHFHYVMSLGAIFAIFAGIYFMPK 433

DB 340 WALGFIFLFTVGGTGLIVLANSLSLIDVLDHTYVVAHFHYVLSMGAVFAIMGFIHWFPL 399

QY 434 FSGRAPPEWAAKLHFWTFFIGANVTFFPQHFLGRQGMPPRYIDYPAFALWNVKVSYGAF 493

DB 400 FSGYTLDDQTYAKIHFTIMFIGNLTFFPQHFLGSGMPRRYSYDPAITWNILSSVGSF 459

QY 494 LAFASLFFIIVFYVTLVAGRRTNPNMGEFAD-----TLEWTLPSPPAHT 541

DB 460 ISLTAVMLMIFMI-----WEAFASKKVLWMBEPPSNMLEWLYGCPPPYHT 504

QY 542 FE 543

DB 505 FE 506

RESULT 3

US-09-252-991A-17309

; Sequence 17309, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 17309

; LENGTH: 425

; TYPE: PRT

```

77  ---EANYHNEVPTHGVIMILFMANPFFAF-GLWVVVPIQLGARDVAFVPMVNVSPMLF 131
138  IAGTANGVASLAPAGDGGQLGSGGVHLYPPLESTR-EAGYSMDLAIFAVHLSGASSINGA 196
132  FAGMILFNLS-PIVGG- ----SPAAGWNTYAPLAGEFSPGPGVNYLIAIQISIGSLMTG 186
197  INMITTFLNMRAPGMTLHKVPLPFSWISFITAMLIILLALPVLACAITMLLTDNRNFGTTFFN 256
187  INFFVTILCKTPTMKFQMOMPHESVTTFTTLLIVILAFPVFTVALALMTADTRIFGTQFTT 246
257  PAGGDPILYQHILWFEGHPEVYIILPGFGIISHVVSTFSKPVPGYLPMTYAMVAIGV 316
247  VANGGMPLWANFFWYMGHPEVYIVILPAGMYSEIPTFARXRLFGHQSMIWATAGIAF 306
317  LGFVVAHMYTVMGSLTQOSYFMLATMVLAVPTG:KIFSWIATMGGSEVEFKSPMLWAF 376
307  LGLFLWVHHFTMGNGALLNSPFSISTMLIGVPTGVKLFNWLITLYKGRITFSPMLFSL 366
377  GRMFLFTVGGVTGIVLQAQGLDRAYHDTYVVVAHFHYVMSLGAIFAIFAGIYFMPKFSG 436
367  APTENFLLLGGVTVGLMASADYQYXNTYFVLVAHFHYTLTVGVVVFACLAGLIWFYPKWVG 426
437  RAPPEWAAKLHWTFTTIGANVTFFPOHFFLGROGMPBRYIDYPFAFALW--NKYSSVYGAEL 494
427  YKINETLWKCWFMMFGVNCVLPQIILGDLGNPRRLTYMPSDCGWLLNFISTIGAVL 486
495  AFASFTFFIVFYVTLVAGRRTRPRNPWGEFADTLEWTLFSP-PPAHTFETLPKRSDWD- 552
487  MAIGPLFLVASIVYSHIKAPREATGDNWDGLGRTLWSTASAIPPKYNFATPDWNDYDT 546
553  -----KHPSH 557
547  FVDMKSHGRH 556

RESULT 5
US-09-134-001C-3430
; Sequence 3430, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHY-
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3430
; LENGTH: 668
; TYPE: PRN
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3430

```

```

Query Match      35.9%; Score 1078.5; DB 3; Length 668;
Best Local Similarity 38.5%; Pred. No. 1.6e-97;
Matches 212; Conservative 102; Mismatches 203; Indels 33; Gaps 8;

Qy      18  FTFWFSTNKHDKGLLYLYAAGVVGFSVLFTVYVMELEMDPGVQVCMLEGARLIADASQ  77
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      50  YKEWFVSVDHKKTGIMYLLISAVLNFVRGGIDALMLRTQLTTPDNKFM-----  96

Qy      78  TCTANGHLNMVYVYTGILMMFFVGIPALEFGGPGNYLMPQLQIGAPDNMAFPMNMNLSEWLF  137
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      97  ----EANHYNVEVTHGVIMIIFNAEPFIF-GLNMVVIPLQLGARDVAFPMNMNVSEWLF  151

Qy     138  IAGTAMGVASLFPAGGDDQLGSGVGWVLYPPLSTR-BAGYSMDLAIFAVHLSGASSINGA  196

Db     152  FAGMILFNLS-FVGG----SPAAGMTNYAPLAGEFSPGPGVNYLLAIATIGTSGSLMTG  206

```

```

; ORGANISM: Pseudomonas aeruginosa
US-09-252-99JA-17309

Query Match 42.6%; Score 1278; DB 4; Length 425;
Best Local Similarity 55.6%; Pred. No. 1.9e-117;
Matches 234; Conservative 71; Mismatches 110; Indels 6; Gaps 3

QY 127 PMNNLSFWLPIAGTAMGVASLPAPGGDGLQSGVGNVLYPPPLUSTREAGYSMDLAIFAVH 186
DB 9 PMNNFSFWLLPAAFGLLVSTLFPGG---GNFNGWTFYAPLSTTTFAPHSVTFFIFAIH 64
QY 187 LSGAGSIKGAINMITFLNMEAPGMWTLHKVPLFSWSFITAWLILALPVLAGAATMLLT 246
DB 65 LAGISIKGAINVATINLEAPGWTLMKXPLFWTWLITAFLLIAWVPLAGVVTMLM 124
QY 247 DRNFGTTFFNPAGGDPILYOHLLWFFGHPEVYIILPGFGIISHVVSTFSKKPVGYLP 306
DB 125 DIHFGTSFSAAGGDPVLFQHVWFVFFGHPEVYIMILPAFGAVSAIIPFARKPLFGVTS 184
QY 307 NYVAMVATGVLFVYVAHHMYTVGMSLTQOSYFPLATMTAVTGIGIKISWATWNGGSV 366
DB 185 MYVATSAFISFVYVAHHMFVVGIPVTGELFFMYATMLIAVTPGVKVFNNWVTWMBGSL 244
QY 367 EFKSPMLWAFGMFLFTVGGVTGIVLAQAGIDRAYHDTYVVAHFHYVMSLGAIFAIFAG 426
DB 245 TETPMLEFAVAVILFTTIGGSGMLAIPADPOYHDTYFVVAHFHYVLVPGAIFGIFAS 304
QY 427 IVFYMPKSGRAFPFWAAKHLHFWTFTFCANVTFFPQHFLGRQGNPRYIDYEPFAFMNK 486
DB 305 AYVWLPKWTGHMYDETLLGKLHFWMSFIGMNLAFPPMHFVGLAGNPRRIPDYNLQFADFN 364
QY 487 VSSYGAFAPASFLFFIVFYVYLVAGRRETRPNPWGEFADTLEWLTPSPPAHHTPETLP 546
DB 365 VSSIGAFMFGTQLLFLFI-VKIRGKGPAPAKPW-DGAEGLEWSIPSPAPYHITFSTPP 422
QY 547 K 547
DB 423 E 423

RESULT 4.
US-09-710-279-1060
; Sequence 1060, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1060
; LENGTH: 648
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-1060

```

Query Match	35.9%;	Score 1078.5;	DB 4;	Length 648;
Best Local Similarity	38.5%;	Pred. NO. 1.5e-97;		
Matches 212; Conservative 102; Mismatches 203; Indels 33; Gaps 8				
Qy	18	FTREWSTNHKDIGLLYLVAAGVGFISVLFTVMRLMDPGVQYVCLEGARLIADAQQ	77	
Dd	30	YKEWFTSVDHKKIGNYLISAVLMFVRGIDALMLRTQLTIPDNKFL-----	76	
Qy	78	TCTANGHLANNVMVTYHGILMFPFGVIGIPALFCGGFGNLYMPQIAGPDMAFFPMNNLSFWLF	137	

Db 457 GFTLNKRGKIRA-FWFWFIQFFVAFMPYILGFMGMTRRLSQINPE-FHPMLVAAGGA 514
Qy 493 FLAPASLFFIVIVYTLVAGR-----RETRNPNWGEFADTLEWLPSPPAHTPETLPK-- 547
Db 515 AL-IALGIVCQIQIYVSIIRDHLNRLDTGDPWG--GRTLEWSSISPAPFFNFAVEPQVK 571
Qy 548 -RSDW 551
Db 572 ARDEW 576

RESULT 8
US-09-711-164-332
; Sequence 332, Application US/09711164
; Patent No. 6589738
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THERAPY
; CURRENT APPLICATION NUMBER: US/09/711.164
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/164415
; PRIOR FILING DATE: 1999-11-9
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 332
; LENGTH: 663
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-711-164-332

Query Match 33.7%; Score 1010.5; DB 4; Length 663;
Best Local Similarity 37.9%; Pred. No. 7.9e-91;
Matches 212; Conservative 104; Mismatches 193; Indels 51; Gaps 14;
Qy 21 WFMSTNHXDI GLLYLVAAGV---GFISVLFTVMRLMELDPGVQVCMLEGARLIADASQ 77
Db 48 WLTSDVHKRLGIMYIIIVMLLRGFADAIM-----MRSQQALASAGE 90
Qy 78 TCTANGHLNMVYVYHGIILMMFFVIGIPALFGGNYLMLPQIGADPMAPPMNLSWFL 137
Db 91 AGFLPPHYDQFTAHGVIMIFFVAMPFVI-GLMNLVVPQLGARDVAFPPFNLSWFT 149
Qy 138 TAGTANGVASLPAPGGDGLGSGVGVVLYPLPLSTRE--AGYSMDLAIFAVHLSGASSIMG 195
Db 150 VVGVLNVNLSL---GVGEFAQ-TGMLAYPLPSGIEYSPGVGVVYVWKSLSGIGTTLT 204
Qy 196 AINMTTFLNMRAPGNTLHKVLPFSWIFITAWLILLALPVLGAIITMLTDRNFGTTF 255
Db 205 GINFFVTILKMRAPGNTMFKMPVFTWASLCANVLIIAGFPPLTIVTVALTLDRLYLGTHFF 264
Qy 256 NPAGGDPILYOHILWFFGHPEVYIIILPGFGIISHVSTFSKPKVPFGLPMVYAMVAIG 315
Db 265 TNDMGGNMMYINLIWAGHPEVYIIILPVGVTFSEIAATFSKELFGYTSILVATVCIT 324
Qy 316 VLGFVVAHMYTVGNSLTQQSYFMLATVIAVPTGKIFSWIATMWGGSVEFKSPMLWA 375
Db 325 VLSFIVLWHPFTMGAGANVNAFFGITTMIIAIPTGVKIFENLFTMYQGRIVFHSAMLT 384
Qy 376 FGMELFTVGGVGTGIVLAQAGLDRAVHDYVYVAHFHYVMSLGATFALFAGIFYMPKFS 435
Db 385 IGFIIVTFSVGGMTGVLLAVPGADFVHNSLFLIAHFHNVIIGGVVFGFAGTYWPKAF 444
Qy 436 GRAFPEWAAKLFHFTFFIGANVTFFPQHFPGRCGMRR---YIDYPEAFALMNKVSYGA 492
Db 445 GFKLNETWGRKRAFVFIIGFFVAFMPLYALGFMGMTRRLSQID-PQ-FHTMLMIAASGA 502
Qy 493 FLAPASLFFIVIVYTLVAGRTRP---NPMGFEADTLEWLPSPPAHTPETLP--- 546
Db 503 VL-IALGILCLVQIYVSIIRDQNRDLTGDPWG--GRTLEWATSSPPFFNFVAVPVH 559

Qy 547 -KRSDW-----KHPSH 557
Db 560 ERDAFWEMKEGEAYKKPDH 579
RESULT 9
US-09-489-039A-13088
; Sequence 13088, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489.039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13088
; LENGTH: 677
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13088

Query Match 33.7%; Score 1010.5; DB 4; Length 677;
Best Local Similarity 37.3%; Pred. No. 8.2e-91;
Matches 210; Conservative 107; Mismatches 195; Indels 51; Gaps 14;
Qy 18 FTFWSTNHKDIGLLYLVAAGV---GFISVLFTVMLELMDPGVQVCMLEGARLIAD 74
Db 59 WNEWLTSDVHKRLGIMYIIIVMLLRGFADAIM-----MRSQQVLAS 101
Qy 75 ASOTCTANGHLNMVYVYHGIILMMFFVIGIPALFGGNYLMLPQIGADPMAPPMNLSF 134
Db 102 AGEAGLPPHYDQFTAHGVIMIFFVAMPFVI-GLMNLVVPQLGARDVAFPPFNLSF 160
Qy 135 WLFIAGTAMCVASLPAPGGDGLGSGVGVVLYPLPLSTRE--AGYSMDLAIFAVHLSGASS 192
Db 161 WFTVGVVILNVLSL---GVGEFAQ-TGMLAYPLPSGIEYSPGVGVVYVWKSLSGIGT 215
Qy 193 IMGAINMTTFLNMRAPGNTLHKVLPFSWIFITAWLILLALPVLGAIITMLTDRNFGT 252
Db 216 TLGINFFVTIILKMRAPGNTMFKMPVFSWASLCANILIIASFPPLTIVTVALTLDRLYGT 275
Qy 253 TFPNAGGDPILYOHILWFFGHPEVYIIILPGFGIISHVSTFSKPKVPFGLPMVYAMV 312
Db 276 HFTNDMGGNMMYINLIWAGHPEVYIIILPVGVTFSEIAATFSKRLFGYTSLVATV 335
Qy 313 AIGVLGFVVAHMYTVGNSLTQQSYFMLATVIAVPTGKIFSWIATMWGGSVEFKSPM 372
Db 336 CITVLSFIVLWHPFTMGAGANVNAFFGITTMIIAIPTGVKIFENLFTMYQGRIVNSAM 395
Qy 373 LWAFGMELFTVGGVGTGIVLAQAGLDRAVHDYVYVAHFHYVMSLGAI FALFAGIFYMP 432
Db 396 MWTIGFIVTFSVGGMTGVLLAVPGADFVHNSLFLIAHFHNVIIGGVVFCFAGLTWWP 455
Qy 433 KFSGRAPPEWAAKLFHFTFFIGANVTFFPQHFPGRCGMRR---YIDYPEAFALMNKVS 489
Db 456 KAFGFTLNETWGRKRAFVFIIGFFVAFMPLYALGFMGMTRRLSQID-PQ-FHPMLVIAA 513
Qy 490 YGAPLAFASLFFIVIVYTLVAGRTRP---NPMGFEADTLEWLPSPPAHTPETLP 546
Db 514 CGAALACILCLVQIYVSIIRDQNRDLTGDPWG--GRTLEWATSSPPFFNFVAVP 570
Qy 547 KRSD-----KHPSH 557
Db 571 QVHERDAFWEMKEGEAYKQPAH 593

RESULT 10
US-09-252-981A-22396
; Sequence 22396, Application US/09252991A

```
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074.788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094.190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22396
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22396

Query Match
32.5%; Score 974.5; DB 4; Length 660;
Best Local Similarity 38.2%; Pred. No. 2.8e-87;
Matches 211; Conservative 97; Mismatches 193; Indels 51; Gaps 15;

QY 18 FTRFWSTNKHDKIGLLYLVAAGVY--GFTISVLTFTVYMRLELMDPGVQVCMLEGARLIAD 74
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 47 WTEMLTSVDHKKIGVMYIVVALVMLYRGFAD---AIMEGQL-----ALAEAGN---- 92

QY 75 ASOTCTANGHL-----NNVMVYTHGILMMFFVGIIPALFGGFGNYLMPLOIGAPDMAPPRVN 130
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 93 -----HGVLPEPHDQIHTAGVMIIFMAMP-FMTGLMNLAVPLQIGARDVAFPFILN 144

QY 131 NLSEFWLFIAGTAGVAVSLPAPGDGQSGGVGVWLYPPLSTREAGVS---NDLALFAVH 186
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 145 SLSEFWLVSAMLVNLSL-----GLGEF-ARTGNVAYPLPS--ELAYSQGVGVYIYIWAQ 197

QY 187 LSGASSIMGAINMITTFLNMRAPGMTLHKVPLFSWISFITAWLILALPVLAGAITMLLT 246
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 198 ISCGMTLLTGINFVTKMRAPGMKLMQPIFTWCTCFANILIVASPPILTAAGLLSL 257

QY 247 DRNFGTTFENPAGGDPILYQHLLWFFGHPEVYIILPGFIISHVSTFSKKPVGYLP 306
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 258 DRYLDHMFNEUGGNAMVINLFWANGHPEVYIILPAGFIPSEVTATFAGKRMFGYKS 317

QY 307 MYVAMVAIGVLGVVVAHMYTVGMSLTQOSYFMLATMVAIVPTGKIPSWIATWGGSV 366
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 318 MYWASAAITFLGFTVWLHFFHTMGSGDVNGFFGVATMLISPTGVKLFNLTIIYKGR 377

QY 367 EFKSPMLWAFGNFLFTVGGVTVGLVLAQGLDRAYHDTYVVAHMYVMSIGAIFAIPAG 426
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 378 RSTPLTLTGLFNMVTTIGMTGVLVLAIPCADFLHNSLFLIAHPHTIIGGAVFGYLAG 437

QY 427 IYFYMFKFSRAPPEWAAKHLFWFTFIFGANVTFFFOHFLGROCMRRYVDYPEAFALW 486
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 438 FARWPKAFGLTDEKWKRESFNCWLVGVYAFMPLYLILGFMGTRRLNHYDN--PLWKP 495

QY 487 ---VSSYGAFAPASPLFIVFYVTLVAGRR---ETRNPWGEFADLTWLTLPSPPAH 540
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 496 YLVYAFFGVAVLFCG-IACQLQLFVSVRNRKQLADVNGDPWE--CRTLEWATSSPPPFY 552

QY 541 TPETLPKRSDWD 552
Db 553 NFAELPKVQDWD 564
```

RESULT 11

```
US-09-248-796A-16350
; Sequence 16350, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
```

```
; CURRENT APPLICATION NUMBER: US/09/248.796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074.725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096.409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16350
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-16350

Query Match
7.8%; Score 234.5; DB 4; Length 98;
Best Local Similarity 42.3%; Pred. No. 1.7e-15;
Matches 47; Conservative 18; Mismatches 31; Indels 15; Gaps 1;

QY 24 STNKHDKIGLLYLVAAGVYGFISVLTFTVYMRLELMDPGVQVCMLEGARLIADASQTCTANG 83
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3 STSHDKIGLLYLIVGMSAVATGMSVILRELSGGPGMFL-----HGNN 47

QY 84 HLMNVMVYTHGILMMFFVGIIPALFGGFGNYLMPLOIGAPDMAPPRMNLSE 134
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 48 QVFNVLVTGHAIAIMFLFVMPILIGSGFNYFLPIMIGAVDMAFARLNNISF 98

RESULT 12
US-09-354-129-2
; Sequence 2, Application US/09354129
; Patent No. 6136588
; GENERAL INFORMATION:
; APPLICANT: YE, RICK W
; TITLE OF INVENTION: GENES ENCODING DENITRIFICATION REACTIONS
; FILE REFERENCE: CL-1255
; CURRENT APPLICATION NUMBER: US/09/354.129
; CURRENT FILING DATE: 1999-07-15
; EARLIER APPLICATION NUMBER: 60/093.181
; EARLIER FILING DATE: July 17, 1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Pseudomonas sp. strain G-179
US-09-354-129-2

Query Match
6.2%; Score 186; DB 3; Length 472;
Best Local Similarity 21.7%; Pred. No. 8.8e-10;
Matches 98; Conservative 77; Mismatches 173; Indels 104; Gaps 23;

QY 89 MVTYHGLMMFFVGIIPALFGGFGNYLMPLOIGAPDMAPPRMNLSEFWLFIAGTAMGVASL 148
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 51 MIHTNALIVWLLIG---FMGATYYLLP-EAEETELYSKLAQAQFWIFLAAIAV-- 103

QY 149 FAPCGDQSGGVGVWLYPPLSTREAGYSMDLAFPAVHLSGASSIMGAINMITTFLNMA 208
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 104 -----GYMFKIHGREFLEQFFIKIGIVIV-----CLMFLFNVMTSLKGRK 146

QY 209 PGMTHLHKVPDLS-WSFITAWLILALPVLAGAITMLLTDNRNFGTTFFNPAG-GGDPILY 266
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 147 TVVT--NILIFGLWGIAIP-----FLFSFYFNPANLADKXVW 181

QY 267 QHI--LWFFGHPEVYIILPGF-----GIISHVSTFSKKPVGVGLPVMYAWAI----- 314
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 182 WYVVLHLMVEGVWELIMASVLAFLMKLNGIDREVVEKW-----LYVIVGLALFS 230

QY 315 GVLGVVVAHMYTVGM-SLTQOSYFMLATMVAIVPTGKIPSWIATMW--GGSVEFKSP 371
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 231 GILG---TGHHYTWIGAPGVWQWIGLSTLEVPDFTVIFVT--WGAARKHPNRAA 286

QY 372 MLWAFG-FMFLFTVGGVTVGLVLAQGLDRAYHDTYVVAHFF-----YVMSIGAIFAIF 424
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 287 LLWSIGSVMAFFGAGVWGLHTLSSVNYTHTGTQVTAAGHGLAFPGAVVMLNLAIMA-- 344
```


Db 286 INGMWLSGAWHKLRDDPILRFLVSLAFYGMSTFEGPMMAIKTVNSLSHYTDWTIGHVH 345
QY 413 YVMSLGAIFAIFAG-IYFYMPKTSGRAFPDWAAKL--HFWTFFIGANVTFFFQHFPLG--R 467
Db 346 -AGALGWANISIGTLYHMIPIKLGRRMHSVGLINAHFWLATIG-TVLYIASMWVNGIT 403
QY 468 QCMPPRYIDYPEAFALWKNKVSSYGAFAPASFLFFIVFYTLVA 512
Db 404 QGL-----MWRVNEEDGT-LTYS-----FVEALVA 427

Search completed: October 18, 2004, 23:27:58
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 18, 2004, 23:25:00 ; Search time 131 Seconds
(without alignments)
1374.722 Million cell updates/sec

Title: US-09-712-768A-2

Perfect score: 3000

Sequence: 1 MADAAIHGHDBHKQGFTR.....PAHTFETLPKRSWDKHPSH 557

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1360919 seqs, 323318874 residues

Total number of hits satisfying chosen parameters: 1360919

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pap:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pap:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pap:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pap:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pap:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pap:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pap:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pap:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pap:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pap:*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap:*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1531	51.0	513	9	US-09-098-079-15
2	1531	51.0	513	15	Sequence 15, Appl
3	1531	51.0	513	15	Sequence 32, Appl
4	1531	51.0	513	17	US-10-428-487-32
5	1531	51.0	513	17	US-10-231-956A-64
6	1531	51.0	513	17	US-10-770-668-72
7	1531	51.0	513	16	US-10-408-765A-191
8	1531	51.0	513	14	US-10-127-032-117
9	1531	51.0	513	14	US-10-389-647-362
10	1531	51.0	513	14	US-10-156-761-14064
11	1531	51.0	513	14	US-10-282-122A-62858
12	1531	51.0	513	14	US-10-282-122A-64832
13	1531	51.0	513	14	US-10-156-761-13575
14	1531	51.0	513	15	US-10-282-122A-63927
15	1531	51.0	513	15	US-10-282-122A-61730
16	1531	51.0	513	15	US-10-282-122A-54166

15	1169.5	39.0	581	9	US-09-738-626-6269	Sequence 6269, Ap
16	1169.5	39.0	584	9	US-09-945-825-2	Sequence 2, Appli
17	1169.5	39.0	584	15	US-10-380-055-2	Sequence 2, Appli
18	1160	38.7	552	16	US-10-781-014-854	Sequence 654, App
19	1138.5	38.0	659	15	US-10-282-122A-60970	Sequence 60970, A
20	1098.5	36.6	644	15	US-10-282-122A-46238	Sequence 46238, A
21	1078.5	35.9	662	15	US-10-282-122A-70569	Sequence 70569, A
22	1060.5	35.4	662	15	US-10-282-122A-44486	Sequence 44486, A
23	1059.5	35.3	649	9	US-09-815-242-5333	Sequence 5333, Ap
24	1059.5	35.3	662	9	US-09-815-242-12343	Sequence 12343, A
25	1056.5	35.2	642	15	US-10-282-122A-51332	Sequence 51332, A
26	1046.5	34.9	648	15	US-10-282-122A-71411	Sequence 71411, A
27	1044	34.8	642	15	US-10-282-122A-44805	Sequence 44805, A
28	1040	34.7	660	15	US-10-282-122A-68962	Sequence 68962, A
29	1023.5	34.1	646	15	US-10-282-122A-50435	Sequence 50435, A
30	1019.5	34.0	663	15	US-10-282-122A-78152	Sequence 78152, A
31	1010.5	33.7	663	9	US-09-815-242-10070	Sequence 10070, A
32	1010.5	33.7	663	14	US-10-287-274-332	Sequence 332, App
33	1010.5	33.7	663	15	US-10-282-122A-56452	Sequence 56452, A
34	1010.5	33.7	663	15	US-10-282-122A-59897	Sequence 59897, A
35	1004.5	33.5	670	15	US-10-282-122A-69771	Sequence 69771, A
36	1004	33.5	668	15	US-10-282-122A-50550	Sequence 50550, A
37	1003	33.4	668	15	US-10-282-122A-47598	Sequence 47598, A
38	997.5	33.2	663	15	US-10-282-122A-75460	Sequence 75460, A
39	997	33.2	662	15	US-10-282-122A-49239	Sequence 49239, A
40	995.5	33.2	672	15	US-10-282-122A-68178	Sequence 68178, A
41	990.5	33.0	663	9	US-09-815-242-11080	Sequence 11080, A
42	970.5	32.4	658	9	US-09-815-242-11824	Sequence 11824, A
43	970.5	32.4	658	15	US-10-282-122A-66271	Sequence 66271, A
44	970.5	32.4	658	15	US-10-389-647-413	Sequence 413, App
45	957	31.9	652	15	US-10-282-122A-73132	Sequence 73132, A

ALIGNMENTS

RESULT 1

US-09-098-079-15
; Sequence 15, Application US/09098079
; Patent No. US20020064773A1
; GENERAL INFORMATION:
; APPLICANT: Heirnstadt, Corrina
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Cleverger, William
; APPLICANT: Fahy, Eoin F.
; APPLICANT: Davis, Robert E.
; TITLE OF INVENTION: DIAGNOSTIC METHOD BASED ON QUANTIFICATION OF
; TITLE OF INVENTION: EXTRAMITOCHONDRIAL DNA
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/098,079
; FILING DATE: 15-JUN-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rosenman Ph.D., Stephen J.
; REGISTRATION NUMBER: 43,058
; REFERENCE/DOCKET NUMBER: 660088.416
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:
 LENGTH: 513 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear

US-09-098-079-15

```

Query Match          51.0%; Score 1531; DB 9; Length 513;
Best Local Similarity 52.4%; Pred. No. 3.1e-126;
Matches 284; Conservative 81; Mismatches 125; Indels 52; Gaps 6;

QY 17 FFRWFMSSTNKHDKIGLLYLVAAGVGVGFIISVLFTVYMELEMDPGVQVCMLEGARLIADAS 76
Db 2 FADRWLFSTNKHDKIGLLYLVAAGVGVGFIISVLFTVYMELEMDPGVQVCMLEGARLIADAS 76
QY 77 QTCTANGHLNVMVYVYHILMMPFVGIPALFGGFGNLYMPLQIGAPDMAPPRMNNLSFWL 136
Db 2 FADRWLFSTNKHDKIGLLYLVAAGVGVGFIISVLFTVYMELEMDPGVQVCMLEGARLIADAS 76
QY 46 -NLLGNDHIIYVITAHAFVMIFFMVPIMIGFGNWLPLMIGAPDMAPPRMNNLSFWL 104
Db 46 -NLLGNDHIIYVITAHAFVMIFFMVPIMIGFGNWLPLMIGAPDMAPPRMNNLSFWL 104
QY 137 FIAGTMGVASLFPAGDGGQLGSGVGVWLYPPPLSTR--EAGYSMDLAIFAVHLSGASSTM 194
Db 105 LPPSLLLLLASAMV-----EAGAGTGTWVYPPPLAGNYSHPGASVDLTIFSLHLAGVSSIL 159
QY 195 GAINMTITFLNMRAPGMLHKVPLFSWSIFITAWLILLALPVLGAIITMLLTDNRFGTTF 254
Db 160 GAINFITTIINMKPPANTQYQTLFVMSVLITAVLLLSLPLVLAAGITMLLTDNRNLNITF 219
QY 255 FNPAGGDPILYQHILWFHFGHPVEVYIILPGFGIISHVSTFS--KXPVFGYLPWVYAMVA 313
Db 220 FDPAGGDPILYQHILWFHFGHPVEVYIILPGFGIISHVSTFS--KXPVFGYLPWVYAMVA 313
QY 314 IGVLFVVAHMYTVGMSLTQOSYFMLATMVIATPTGKIFSWIATMNGSVBEKSPML 373
Db 280 IGVLFVVAHMYTVGMSLTQOSYFMLATMVIATPTGKIFSWIATMNGSVBEKSPML 373
QY 374 WAFGFMFLFTVGGTGVILQAAGLDRAVHDTYVVAHFVYVMSLGAIFAFAFIAGIYFMPK 433
Db 340 WALGFIFLFTVGGTGVILQAAGLDRAVHDTYVVAHFVYVMSLGAIFAFAFIAGIYFMPK 433
QY 434 FSGRAPEWAAKLHFTFFIGANVTFFPQHFLGRQGMPPRYIDYPEAFALMKNVSSYGAF 493
Db 400 FSGYTLDDQYAKHFTFFIGANVTFFPQHFLGRQGMPPRYIDYPEAFALMKNVSSYGAF 493
QY 494 LAFASFLFFIVFYVTLVAGRRTRPNWGEFAD-----TLEWTLPSPPPAHT 541
Db 460 ISLTAVNLMIFMI-----WEAFASKRKLVMVEEPMNLEWLYGCPPPYHT 504
QY 542 FE 543
Db 505 FE 506
  
```

RESULT 2
 US-10-428-487-32
 ; Sequence 32, Application US/10428487
 ; Publication No. US20040006780A1
 ; GENERAL INFORMATION:
 ; APPLICANT: RASTELLI, LUCA K.
 ; APPLICANT: GERBER, HANS-PETER
 ; TITLE OF INVENTION: VEGF-MODULATED GENES AND METHODS EMPLOYING THEM
 ; FILE REFERENCE: 09800080-0103
 ; CURRENT APPLICATION NUMBER: US/10/428,487
 ; PRIOR FILING DATE: 2003-05-02
 ; PRIOR APPLICATION NUMBER: 09/815,153
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,201
 ; NUMBER OF SEQ ID NOS: 84
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 32
 ; LENGTH: 513
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens

```

US-10-428-487-32

Query Match          51.0%; Score 1531; DB 15; Length 513;
Best Local Similarity 52.4%; Pred. No. 3.1e-126;
Matches 284; Conservative 81; Mismatches 125; Indels 52; Gaps 6;

QY 17 FFRWFMSSTNKHDKIGLLYLVAAGVGVGFIISVLFTVYMELEMDPGVQVCMLEGARLIADAS 76
Db 2 FADRWLFSTNKHDKIGLLYLVAAGVGVGFIISVLFTVYMELEMDPGVQVCMLEGARLIADAS 76
QY 77 QTCTANGHLNVMVYVYHILMMPFVGIPALFGGFGNLYMPLQIGAPDMAPPRMNNLSFWL 136
Db 46 -NLLGNDHIIYVITAHAFVMIFFMVPIMIGFGNWLPLMIGAPDMAPPRMNNLSFWL 104
QY 137 FIAGTMGVASLFPAGDGGQLGSGVGVWLYPPPLSTR--EAGYSMDLAIFAVHLSGASSTM 194
Db 105 LPPSLLLLLASAMV-----EAGAGTGTWVYPPPLAGNYSHPGASVDLTIFSLHLAGVSSIL 159
QY 195 GAINMTITFLNMRAPGMLHKVPLFSWSIFITAWLILLALPVLGAIITMLLTDNRFGTTF 254
Db 160 GAINFITTIINMKPPANTQYQTLFVMSVLITAVLLLSLPLVLAAGITMLLTDNRNLNITF 219
QY 255 FNPAGGDPILYQHILWFHFGHPVEVYIILPGFGIISHVSTFS--KXPVFGYLPWVYAMVA 313
Db 220 FDPAGGDPILYQHILWFHFGHPVEVYIILPGFGIISHVSTFS--KXPVFGYLPWVYAMVA 313
QY 314 IGVLFVVAHMYTVGMSLTQOSYFMLATMVIATPTGKIFSWIATMNGSVBEKSPML 373
Db 280 IGVLFVVAHMYTVGMSLTQOSYFMLATMVIATPTGKIFSWIATMNGSVBEKSPML 373
QY 374 WAFGFMFLFTVGGTGVILQAAGLDRAVHDTYVVAHFVYVMSLGAIFAFAFIAGIYFMPK 433
Db 340 WALGFIFLFTVGGTGVILQAAGLDRAVHDTYVVAHFVYVMSLGAIFAFAFIAGIYFMPK 433
QY 434 FSGRAPEWAAKLHFTFFIGANVTFFPQHFLGRQGMPPRYIDYPEAFALMKNVSSYGAF 493
Db 400 FSGYTLDDQYAKHFTFFIGANVTFFPQHFLGRQGMPPRYIDYPEAFALMKNVSSYGAF 493
QY 494 LAFASFLFFIVFYVTLVAGRRTRPNWGEFAD-----TLEWTLPSPPPAHT 541
Db 460 ISLTAVNLMIFMI-----WEAFASKRKLVMVEEPMNLEWLYGCPPPYHT 504
QY 542 FE 543
Db 505 FE 506
  
```

RESULT 3
 US-10-231-956A-64
 ; Sequence 64, Application US/10231956A
 ; Publication No. US20040053233A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lorens, James B.
 ; APPLICANT: Xu, Weiduan
 ; APPLICANT: Bogenberger, Jakob
 ; APPLICANT: Holland, Sacha
 ; APPLICANT: Rigol Pharmaceuticals, Incorporated
 ; TITLE OF INVENTION: Modulators of Angiogenesis
 ; FILE REFERENCE: 021044-004100US
 ; CURRENT APPLICATION NUMBER: US/10/231,956A
 ; CURRENT FILING DATE: 2001-08-30
 ; NUMBER OF SEQ ID NOS: 522
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 64
 ; LENGTH: 513
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-231-956A-64

Query Match
 Best Local Similarity 52.4%; Pred. No. 3.1e-126;
 Matches 284; Conservative 81; Mismatches 125; Indels 52; Gaps 6;

```
QY 17 FTRWFMSTNNHKDGLGLLYLVAAGVGVGFISVLFTVYMELELMDPGVQVMCLGEGARLIADAS 76
Db 2 FADRWLFSTNNHKDGLGLLYLVAAGVGVGFISVLFTVYMELELMDPGVQVMCLGEGARLIADAS 45
QY 77 QTCFANGHLNNVMTVYHGIILMMFFVGPALFGGFGNVLMPLOIGAPDMAPFRNNNSFWL 136
Db 46 -NLLGNDHIYNNIVTAHAFVWIFWVMPIMIGFGNVLMPLOIGAPDMAPFRNNNSFWL 104
QY 137 FIAGTMGVASLFAFGDGGQGLGSGVGVLYPLSTR--EAGYSMDLAIFAVHLSGASSIM 194
Db 105 LPPSLLLLLASAMV-----EAGAGTGTVPYPLAGNYSHPGASVDLTIFSLHLAGVSSIL 159
QY 195 GAINMITTFLNMRAPGMLTKHVPFLFSWISFITTAWLILALPVLGAIITMLLTDNRNFGTTF 254
Db 160 GAINFITTIINKKPPAMTQYQTPFLFVWSVLITAVLLLSLPLAAGITMLLTDNRNLNTTF 219
QY 255 FNPAGGGDPILYOHILWFFGHPEVYIIILPGFGIISHVVSTFS--KKPVFGVLPVYAMVA 313
Db 220 FDPAGGGDPILYOHILWFFGHPEVYIIILPGFGIISHVVSTFS--KKPVFGVLPVYAMVA 279
QY 314 IGVLFVVAHMYTGVMSLTQOOSVFMATWIAVPTGIKIFSWIATMWGGSVEFKSPML 373
Db 280 IGVLFVVAHMYTGVMSLTQOOSVFMATWIAVPTGIKIFSWIATMWGGSVEFKSPML 339
QY 374 WAFGFMELFTVGGTGVIGVLAQGLDRAYHDTYVVAHPHYVMSLGAIFAIFAGIYFYMFK 433
Db 340 WALGFILFTVGGTGVIGVLAQGLDRAYHDTYVVAHPHYVMSLGAIFAIFAGIYFYMFK 399
QY 434 FSGRAFPPEWAAKLHFWTFFIGANVTFFPOHFLGQGMPPRYIDYPEAFALMNKVSYSYCAF 493
Db 400 FSGYTLDTQYAKIHFTIMEFIGNLTFFPOHFLGQGMPPRYIDYPEAFALMNKVSYSYCAF 459
QY 494 LAFASFLFFIVFVYTLVAGRRETRPNWGEAD-----TLEWTLSPSPPAHT 541
Db 460 ISLTAVMLMIFMI-----WEAFASKRKLVMVEEPSMNLWLYGCCPPPYHT 504
QY 542 FE 543
Db 505 FE 506

RESULT 4
US-10-770-668-72
; Sequence 72, Application US/10770668
; Publication No. US20040191843A1
; GENERAL INFORMATION:
; APPLICANT: Wright, Susan C.
; APPLICANT: Larrick, James W.
; APPLICANT: Nock, Steffen R.
; APPLICANT: Wilson, David S.
; TITLE OF INVENTION: Cell-Killing Molecules and Methods of Use Thereof
; FILE REFERENCE: ABSALUS-08602
; CURRENT APPLICATION NUMBER: US/10/770,668
; CURRENT FILING DATE: 2004-02-02
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 72
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-770-668-72

Query Match 51.0%; Score 1531; DB 17; Length 513;
Best Local Similarity 52.4%; Pred. No. 3.1e-126;
Matches 284; Conservative 81; Mismatches 125; Indels 52; Gaps 6;

QY 17 FTRWFMSTNNHKDGLGLLYLVAAGVGVGFISVLFTVYMELELMDPGVQVMCLGEGARLIADAS 76
Db 2 FADRWLFSTNNHKDGLGLLYLVAAGVGVGFISVLFTVYMELELMDPGVQVMCLGEGARLIADAS 45
QY 77 QTCFANGHLNNVMTVYHGIILMMFFVGPALFGGFGNVLMPLOIGAPDMAPFRNNNSFWL 136
Db 46 -NLLGNDHIYNNIVTAHAFVWIFWVMPIMIGFGNVLMPLOIGAPDMAPFRNNNSFWL 104
```

```
QY 137 FIAGTMGVASLFAFGDGGQGLGSGVGVLYPLSTR--EAGYSMDLAIFAVHLSGASSIM 194
Db 105 LPPSLLLLLASAMV-----EAGAGTGTVPYPLAGNYSHPGASVDLTIFSLHLAGVSSIL 159
QY 195 GAINMITTFLNMRAPGMLTKHVPFLFSWISFITTAWLILALPVLGAIITMLLTDNRNFGTTF 254
Db 160 GAINFITTIINKKPPAMTQYQTPFLFVWSVLITAVLLLSLPLAAGITMLLTDNRNLNTTF 219
QY 255 FNPAGGGDPILYOHILWFFGHPEVYIIILPGFGIISHVVSTFS--KKPVFGVLPVYAMVA 313
Db 220 FDPAGGGDPILYOHILWFFGHPEVYIIILPGFGIISHVVSTFS--KKPVFGVLPVYAMVA 279
QY 314 IGVLFVVAHMYTGVMSLTQOOSVFMATWIAVPTGIKIFSWIATMWGGSVEFKSPML 373
Db 280 IGVLFVVAHMYTGVMSLTQOOSVFMATWIAVPTGIKIFSWIATMWGGSVEFKSPML 339
QY 374 WAFGFMELFTVGGTGVIGVLAQGLDRAYHDTYVVAHPHYVMSLGAIFAIFAGIYFYMFK 433
Db 340 WALGFILFTVGGTGVIGVLAQGLDRAYHDTYVVAHPHYVMSLGAIFAIFAGIYFYMFK 399
QY 434 FSGRAFPPEWAAKLHFWTFFIGANVTFFPOHFLGQGMPPRYIDYPEAFALMNKVSYSYCAF 493
Db 400 FSGYTLDTQYAKIHFTIMEFIGNLTFFPOHFLGQGMPPRYIDYPEAFALMNKVSYSYCAF 459
QY 494 LAFASFLFFIVFVYTLVAGRRETRPNWGEAD-----TLEWTLSPSPPAHT 541
Db 460 ISLTAVMLMIFMI-----WEAFASKRKLVMVEEPSMNLWLYGCCPPPYHT 504
QY 542 FE 543
Db 505 FE 506

RESULT 5
US-10-408-765A-191
; Sequence 191, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fany, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Marnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 191
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-191

Query Match 50.9%; Score 1528; DB 16; Length 513;
Best Local Similarity 52.4%; Pred. No. 5.7e-126;
Matches 284; Conservative 81; Mismatches 125; Indels 52; Gaps 6;

QY 17 FTRWFMSTNNHKDGLGLLYLVAAGVGVGFISVLFTVYMELELMDPGVQVMCLGEGARLIADAS 76
Db 2 FADRWLFSTNNHKDGLGLLYLVAAGVGVGFISVLFTVYMELELMDPGVQVMCLGEGARLIADAS 45
QY 77 QTCFANGHLNNVMTVYHGIILMMFFVGPALFGGFGNVLMPLOIGAPDMAPFRNNNSFWL 136
Db 46 -NLLGNDHIYNNIVTAHAFVWIFWVMPIMIGFGNVLMPLOIGAPDMAPFRNNNSFWL 104
QY 137 FIAGTMGVASLFAFGDGGQGLGSGVGVLYPLSTR--EAGYSMDLAIFAVHLSGASSIM 194
```

Db 105 LPPSLLLLASAMV-----EAGAGTGWVYPPPLAGNVSHPGASVDLTIESLHLAGVSSIL 159
 QY 195 GAINMTTFLNMRAPGMLHKVPLFWSIFITAWLILALPVLAGAITMLLTDNRPGITP 254
 Db 160 GAINFTIINMRKPPAMTQOTPLFVNSVLITAVALLLSLPLAAGITMLLTDNRNLITP 219
 QY 255 FNPAGGDPILYQHILWFFGHPEVYIILPGFGIISHVYSTFS-KKPFVFGYLPYVAMVA 313
 Db 220 FDPAGGDPILYQHLSWFFGHPEVYIILPGFGWISHVYISGKKEPGYGMWAMMS 279
 QY 314 IGVLFVVAHMYTVGMSLTQOSYFMLATWYAVPTGKIFSWIATWVGVSVEKPSML 373
 Db 280 IGLGFIWAHMYTVGMDVDTREYTSATMIITPTGVKFSWLTATLGSNMKMSAAVL 339
 QY 374 WAFGFMFLFVGGVTGLVLAOAGLDRAHDYTYVVAHFYVMSLGAIRFAIFAGIYFMPK 433
 Db 340 WALGFILFVGGTLGLVLANSSLDIVLHDYTYVVAHFYVMSLGAIRFAIFAGIYFMPK 399
 QY 434 FSGRAPFENAAKHFMTFFIGANVTFFPOHFLGRQGMPPRYIDYDPAFALMNKVSYSYGF 493
 Db 400 FSGYTLDOTYAKIHFAIMFIGNLTFFPOHFLGLSCMPRRYSYDYPDAYTWNILSSVGSF 459
 QY 494 LAFASPLFFVIYVYTLVAGRRTRPNKGEFAD-----TLEWTLPSPPPAHT 541
 Db 460 ISTAVNLMIFMI-----WEAFASKRKVLWVEEPSMNLEWLYGCCPPPYHT 504
 QY 542 FE 543
 Db 505 FE 506

RESULT 6

US-10-127-032-117
 ; Sequence 117, Application US/10127032
 ; Publication No. US20030113742A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Whiteley, Marvin
 ; APPLICANT: Banger, M. Gita
 ; APPLICANT: Lory, Stephen
 ; APPLICANT: Greenberg, Everett Peter
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF
 ; TITLE OF INVENTION: BIOFILM FORMATION
 ; FILE REFERENCE: UIZ-070CP
 ; CURRENT APPLICATION NUMBER: US/10/127,032
 ; CURRENT FILING DATE: 2002-04-19
 ; PRIOR APPLICATION NUMBER: US 60/285,190
 ; PRIOR FILING DATE: 2001-04-20
 ; PRIOR APPLICATION NUMBER: US 60/344,142
 ; PRIOR FILING DATE: 2001-10-24
 ; NUMBER OF SEQ ID NOS: 170
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 117
 ; LENGTH: 530
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-10-127-032-117

Query Match 49.0%; Score 1469.5; DB 14; Length 530;
 Best Local Similarity 51.4%; Pred. No. 8.6e-121;
 Matches 279; Conservative 91; Mismatches 146; Indels 27; Gaps 8;
 QY 7 HGHDDHHE--KQGFTRWFNSTHDKDIGLLYLVAAAGVGVFISVLFTVYVMELEMDPGVOYM 64
 Db 11 HAGDHHGHPAKGLM-RWLTTHNKHDKIGTLYLWFSFMFLGGSMAMVIRAEILFQGLQ-- 67
 QY 65 CUEGARLIADASQTCANGHLNMVYTHGILMMFVGVIPALFGCGVNYLPLQIGAPDM 124
 Db 68 -----IVPEA-----FFNQMTTHGLINMVGAVMPA-FVGLANWMLPLMIGAPDM 111
 QY 125 AFPRMNNLSFWLFIAGTANGVASLPAFGDGGQGLSGVGWVLPPLSTREAGYSMDLAIFA 184
 Db 112 ALPRMNNFSWLLPAAFGLLVSTLFWPGG---GNFGWTFYAPLSTTFAPHSVTFEFA 167
 QY 185 VHLSSASSINGAINMTTFLNMRAPGMLHKVPLFWSIFITAWLILALPVLAGAITML 244
 Db 168 IHLAGISSINGAINMTTFLNMRAPGMLHKVPLFWSIFITAWLILALPVLAGAITML 227

QY 185 VHLSSASSINGAINMTTFLNMRAPGMLHKVPLFWSIFITAWLILALPVLAGAITML 244
 Db 168 IHLAGISSINGAINMTTFLNMRAPGMLHKVPLFWSIFITAWLILALPVLAGAITML 227
 QY 245 LTRNFGTTFNPNAGGDPILYQHILWFFGHPEVYIILPGFGIISHVYSTFSKPKVFGY 304
 Db 228 LMDTHFGTSFSAAGGDPVLFGHVWFVGHPEVYIMILPAFGAVSAIIPTPAKPLFGY 287
 QY 305 LPMYVAMVAGVLCFVVAHMYTVGMSLTQOSYFMLATWYAVPTGKIFSWIATWVG 364
 Db 288 TSMYATASAFUSFVVAHMYTVGIPVTEGELFFMYATMLIAVPTGVKVFNNVITWEG 347
 QY 365 SVEFKSMLWAFGFMFLFTVGGVTGIVLAOAGLDRAHDYTYVVAHFYVMSLGAIRFAIF 424
 Db 348 SLTETPMLFAVAVILFTTGGFSGMLAIPADFOVHDYTYVVAHFYVLPVGAIFGIF 407
 QY 425 AGTYFNPKSGRAFPPEWAAKLFHWTFFIGANVTFFPOHFLGRQGMPPRYIDYDPAFALW 484
 Db 408 ASAYYWLPKWTGHMYDETGLKLFHWSFIGMNLAFPFHMFVGLAGMPRRIPDYNLOFADF 467
 QY 485 NKVSSYGAFLAFASFLFFIVFVYTLVAGRRTRPNKGEFADTLEWTLPSPPPAHTFET 544
 Db 468 NMVSSIGAFMFGTITQLLFLFI-VIKCIRGKGPAPAKPW-DCAGLEWSIPSPAPYHTFET 525
 QY 545 LPK 547
 Db 526 PPE 528

RESULT 7

US-10-389-647-362
 ; Sequence 362, Application US/10389647
 ; Publication No. US20040033549A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GREENBERG, E. Peter
 ; APPLICANT: SCHUSTER, Martin
 ; APPLICANT: LOSTROH, Candi
 ; TITLE OF INVENTION: QUORUM SENSING SIGNALING IN BACTERIA
 ; FILE REFERENCE: UIZ-038CP
 ; CURRENT APPLICATION NUMBER: US/10/389,647
 ; CURRENT FILING DATE: 2003-03-14
 ; PRIOR APPLICATION NUMBER: 09/653730
 ; PRIOR FILING DATE: 2000-09-01
 ; PRIOR APPLICATION NUMBER: 60/153022
 ; PRIOR FILING DATE: 1999-09-03
 ; NUMBER OF SEQ ID NOS: 710
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 362
 ; LENGTH: 530
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-10-389-647-362

Query Match 49.0%; Score 1469.5; DB 15; Length 530;
 Best Local Similarity 51.4%; Pred. No. 8.6e-121;
 Matches 279; Conservative 91; Mismatches 146; Indels 27; Gaps 8;

QY 7 HGHDDHHE--KQGFTRWFNSTHDKDIGLLYLVAAAGVGVFISVLFTVYVMELEMDPGVOYM 64
 Db 11 HAGDHHGHPAKGLM-RWLTTHNKHDKIGTLYLWFSFMFLGGSMAMVIRAEILFQGLQ-- 67
 QY 65 CUEGARLIADASQTCANGHLNMVYTHGILMMFVGVIPALFGCGVNYLPLQIGAPDM 124
 Db 68 -----IVPEA-----FFNQMTTHGLINMVGAVMPA-FVGLANWMLPLMIGAPDM 111
 QY 125 AFPRMNNLSFWLFIAGTANGVASLPAFGDGGQGLSGVGWVLPPLSTREAGYSMDLAIFA 184
 Db 112 ALPRMNNFSWLLPAAFGLLVSTLFWPGG---GNFGWTFYAPLSTTFAPHSVTFEFA 167
 QY 185 VHLSSASSINGAINMTTFLNMRAPGMLHKVPLFWSIFITAWLILALPVLAGAITML 244
 Db 168 IHLAGISSINGAINMTTFLNMRAPGMLHKVPLFWSIFITAWLILALPVLAGAITML 227

Qy	245	LTDNRGCTTFNPAGGDPILYOHILMPFGHPPEVYIIILPGFCILGHVVYSTESKKEPVGY	304
Db	228	LMDIHFGTSFFSAAGGDPVLVQHFVFFGHPPEVYIIMLPAQVSAIIPITFARPLFGY	287
Qy	305	LPNMVYMAIGVLGFEVVAHHMYTVGMSLTQOQSYFMLATVIAVPTGKILPFSWIAVMGG	364
Db	288	TSMWZATASIAFLSPFVVAHHMFVGVIPVTGELFFMYATMLIAVPTGVKVFVNWVTTMEG	347
Qy	365	SVFEPKSBMLWAGFMFLFVCGVTVGLVLAQGLDRAVHDYVYVAHFVHYVMSLGAIFALF	424
Db	348	SLTTFETMLFAVAFVILTTIGFSGSLMLATAPADFVHDDYFVVAHFHVILVPGAIFGIF	407
Qy	425	AGIYFMPKFSGRAPFEWAALKHFWTTFPIGANTWTFPHQLGRCQMPPRYIYDPEAPALW	484
Db	408	ASAYVWLPKWTGHMYDETLGLKHLHFMSPFGIMNLAFPPHMFVGLAGMPRRIPDYNLQFADF	467
Qy	485	NKVSVSGAFAPAFSLFPIVIFVYTVLAGRETRPNPWEGFADTLBWTLLPSPPPAHTPET	544
Db	468	NVUSIGAPFWFTTQLLFLFI-VIKCRGGKPAKAPKW-DGAGLBSWSPSPAPYHTTST	525
Qy	545	LPK	547
Db	526	PPE	528

RESULT 8

```

US-10-156-761-14064, Application US/10156761
; Sequence 14064, Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATORI, MASAHISA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDE
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-20403
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-27536
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14064
; LENGTH: 563
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14064

```

Query Match	41.1%	Score 1234.5	DB 14	Length 563
Best Local Similarity	45.1%	Pred. No. 4.8e-100		
Matches 246	Conservative	91	Mismatches 174	Indels 35
Gaps				12
QY	21	WFMSTNHKDIGLLYLVAAGVGCFVSLEPTVYMRLELMDPGVQYMCLEGARLIADASQTC	80	
DB	25	WLTITDHHKXIGHLYLITSPAFPLIGGVWALVYRAELARPGWQ-----IVDNNQ----	72	
QY	81	ANGHLNNVWVTYHGILMMFFVCI PALFCGPGNYLMPLOIGAPDYAPPRMNNLSFWLFIAG	140	
DB	73	-----FNQLFTLHGITMLLEFATPT-FAGPANE-IMPLOIGSDVAPPRMLNNUSYWLFLFG	126	
QY	141	TAMGVASLPAFGDGGQLSGGVGWLYPLST--REAGYSMDLAIFAVHLSSGSSIMGAIN	198	
DB	127	GLIVLGS LAVSPGSP---AAGCWFAFAYFLNLSRSPGIGADMWIMGLALAGFCTILGSVN	182	
QY	199	MITTFNLNRPAGMTLHKVPLFNSWFTFIWAILLALPVLAGAITMLLTDRNPGCTTFENFA	258	
DB	183	FLTTITGMRAPGTMFRMPDFTWNTLTFTSLIVLMAFPVLAAALVLVSADRFRFGSQVDFDA	242	
QY	259	GGGDRIYLQHIWLFFGHGHPVYIIILPGGFIISHVVVTSFKKPEGVLYPMVYAMVAIVGLZ	318	

```

Db      243  NGG -ALLWOHLFWFFGHPHEVVIIALPPFGGIITBIIPVSRKPIFGYLTIGATAAITGLS 301
Qy      319  FVVAWHNNYTVQMSLTQOSYEMLATMVIAPVTGKIGKIFSMTATWGGSVFEKSPMLWAFQF 378
Db      302  VVVVAHHNFAICAVLL--PPFSFNSFLIAPVTKFNFNTGTMLKGSLSFEIPMLWATGF 359
Qy      379  MELFTVGVGVTSIVIAQAGLDRAVHDTIVVVAHHYHMYMSLGAIPAPAGIYFYMPKESGRA 438
Db      360  LYSFLFGGLTGVIILASPLDPFHVDSFVVAHHYHVVFGTVFVATPGGYFYFWMPKPTGKM 419
Qy      439  FPEWAAKLHFWTFITIGANVTFFPOHFTLGROGMPRRYIDY--PEAFALMNKVSYSYGAFLAF 496
Db      420  LDERLGKIHFWTLFVGPHFTTLVQHWLGAGMPRRYADYLAADGFTALNTLSTICAFLLG 479
Qy      497  ASFLPFIIVYVTLVAGSRETRNPWGEADTLEWTLPSPPPAHTPETLPPK-RSD--WD 552
Db      480  MGLTDFLXNVNWKYARKYGRKVEVDPPWG-FORSLEWTSQPPFHNFTVILPVRVSSPAED 538
Qy      553  -KPSH 557
Db      539  LHHPAH 544

```

RESULT 9

```

US-10-282-122A-62858
; Sequence 62858, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA-034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIORITY FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/369,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62858
; LENGTH: 573
; TYPE: PRT
; ORGANISM: Mycobacterium bovis
US-10-282-122A-62858

```


RESULT 11
US-10-156-761-13575
; Sequence 13575, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION.

Db 369 FLGGTGLVLLASPPDHFVTDSDYFVVAHHYVLCFTIVFSIFAGIYFWFPKWTGELLDE 428
Qy 442 WAALKHPTWTFIGANVTFFQHFGRQGMRRYIDY--PEAFALMNKVSYSYGAFLAFASF 499
Db 429 QLGKLFHFWLTFIGFHTTFLVQHVLGDMGMPRRYADYLPDGTGFCGLNVVSTIGSFILGASM 488
Qy 500 LFTFIVFVYTLVAGRRETRPNWGEFADTLWLPSPPPAHFTETLPK---RSDWDKHP 555
Db 489 FPFVWNVKSWRYGEVVTVDPPWG-YGNSLEWATSCPPRRHNTFELPRISRPAFELHY 547
Qy 556 SH 557
Db 548 PH 549

RESULT 13

US-10-282-122A-61730
; Sequence 61730, Application US/10282122A
; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/369,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 61730

; LENGTH: 543

; TYPE: PRF

; ORGANISM: Mycobacterium avium

US-10-282-122A-61730

Query Match 39.8%; Score 1187; DB 15; Length 543;

Best Local Similarity 43.8%; Pred. No. 7e-96;

Matches 235; Conservative 86; Mismatches 181; Indels 34; Gaps 9;

Qy 30 IGLYLYVAAGVGVFISVLFVYVYVRLMDPGVQVYMCLEGARLIADASQCTTANGHLWNVM 89

Db 2 IGIYIVTCEAFFICGLVALLNRTLEAAPGLQFL-----SNEQFNQL 44

Qy 90 VTYHILMMPFVGIPALFGGNGVYLMPLQIGAPDMAPRNNLSFWLFIAGTANGVASLF 149
Db 45 FTMEGTMLLLYATPVVF--GFANLVLPQIGAPVAFPRNLSFWLFLFGGLAASGFI 103
Qy 150 APGGDGLGSGVGVVLPPLS--TREAGYSMDLAIFAVHLSGASSINGAINMITTFINMR 207
Db 104 VPGG---AADFGWTATPLSDAVHSPGAGGLWITGLIVAGLGTILGAVNMITTVVCMR 159
Qy 208 APGMTLHKVPLFSWISIFITAWLILALPVLGAIITMLLTDNFTFFNPGAGSDPILYQ 267
Db 160 APGMTFRMPIFTWNILVTILILIAFPILTAALFGLAADRHLGAHVYDAANG--VLLWQ 218
Qy 268 HILWFFGHPEVYIIILPGFGIISHVSTFSKVPFGYLPVMYKVAIGVLGVVVAHHMY 327
Db 219 HLFWFFGHPEVYIIALPFGIITEIIPVFARKPVFGYTTILVYATLSAALSVAVWAHMF 278
Qy 328 TVGMSLTQQSYFMATWIVAPFGIKIFSIAIATWGGSVFESKSPMLWAFGFMELFTVGVV 387
Db 279 ATGAVLL--PPFSFMTYLIAPVPTGIKFFNIGTMKGLTFTETPMLFCVGLLTLGLGL 336
Qy 388 TGIVLAQAGLDRAVHDYVYVAHFFHYVMSLGAIFAIFAGIYFYMPKFSGRAFFPEWAAKLH 447
Db 337 TGVMLASPPDHFVTDYFVVAHFFHYVLCFTIVFSIFAGIYFWFPKWTGELLDE 428
Qy 448 FTWTFIGANVTFFQHFGRQGMRRYIDY--PEAFALMNKVSYSYGAFLAFASF 499
Db 397 FWLTFIGFHTTFLVQHVLGDMGMPRRYADYLPDGTGFCGLNVVSTIGSFILGASM 488
Qy 506 FVYTLVAGRRETRPNWGEFADTLWLPSPPPAHFTETLPK---RSDWDKHP 557
Db 457 VFKSWRYGEVVTVDPPWG-YGNSLEWATSCPPRRHNTFELPRISRPAFELHY 511

RESULT 14

US-10-282-122A-54166

; Sequence 54166, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent version 3.1
; SEQ ID NO 54166
; LENGTH: 564
; TYPE: PRT
; ORGANISM: Corynebacterium diptheriae
US-10-282-122A-54166

Query Match 39.2%; Score 1175; DB 15; Length 564;
Best Local Similarity 41.1%; Pred. No. 8.4e-95;
Matches 225; Conservative 102; Mismatches 184; Indels 36; Gaps 10;

Qy 21 WPM--STNHKDIGLYLVAAAGVGVGIVSVLFTVYMELELMDPGVQVMCLGARGLIADASQT 78
Db 30 WXMLTTHDKLGMVYVSWFVFGGLMALLIRAEAFSPGLQFL----- 75
Qy 79 CTANGHLNVMVTVYHILMFFVGPALFPGGNGVYLMPLQIGADPMAPRNNLSFWLFI 138
Db 76 ---SNEQNFQLFTLHGTIMLLAFGTPEVVM-GFSNYILQIGADPADVAPRLNAFGFWITQ 131
Qy 139 ACTANGVASLFPAGGQGLGSGVGVLYPPL--STREAGYSMDLAFIAFVHL-SGASSINGA 196
Db 132 IGVAWKLAFGLTPGG----AADFGWTMLPLADSIHSPGVGGDFWIIIGVATGVTIIASA 187
Qy 197 INMITTFNMRAPGMLTKHKVPLFSWISITAWLIIALLPVLGAIITMLTDNRNFTTFN 256
Db 188 VNMITTLNMRAPGMLTFMFPICWNIQVSVIVLIIPLLTAAALGVWYDRKLGCHYD 247
Qy 257 PAGGDPILYQHILMFFGHPEVYIIILPGFIISHVSTFSKPKVGYLPMYAMVAICV 316
Db 248 PGNNGG-AILWQHLPFFGHPEVYIILPFGFVSEVPIPVFARKPMFGVIGLVFATLSIGM 306
Qy 317 LGFVVAHMYTVGNSLTQCSYFMLATMVIAPVTGKIFSWIATMGGSVFSPMLWAF 376
Db 307 LSMVVAHMYTVTGALL--PFESFWTFLISVPTGVKFNWLGTMWRGHISWEITWTM 364
Qy 377 GMEFLTVGGTVGIVLAQGLDRAYHDTYVVAHMYVNSLGAIFAIFAGIYFYNPKFSG 436
Db 365 GFLVTLFGLGAGIMLASPPDLPHISDTYFVVAHFHYTLFGTVVPASVAGVYFWFPKMTG 424
Qy 437 RAPPENAAKLHFWTPFPGANVTFFPOHFLGRCMPRRYIDYPEA--FALNKKVSSYGAF 494
Db 425 RMLDERLKGHWITFVGPHGTFVQHWVGNMGPRIADYLESQGTTLNQISIVFSL 484
Qy 495 AFASFLFFIVIFYTLVAGERETRPNMGFEADTLEWTLPSPPPAHTFETL 550
Db 485 LGVSVIPFVNNVFKSVRYGIEIVTVDDPWG-YGNSLEWATSCPPPRHNFSLPRISRPA 543
Qy 551 WDKHPSH 557
Db 544 FELHYPH 550

RESULT 15
US-09-738-626-6269
; Sequence 6269, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125

; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patent in ver. 3.0
; SEQ ID NO 6269
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6269

Query Match 39.0%; Score 1169.5; DB 9; Length 581;
Best Local Similarity 39.7%; Pred. No. 2.7e-94;
Matches 221; Conservative 110; Mismatches 188; Indels 37; Gaps 11;

Qy 12 HEKQGFTRFM--STNHKDIGLYLVAAAGVGVGIVSVLFTVYMELELMDPGVQVMCLG 69
Db 19 HARKG-SKAWLMTTTHDKOLGINYIIMSFPFLGGLMALLIRAEFLTGLQFL----- 72
Qy 70 RLIIADASQCTCTANGHLNVMVTVYHILMFFVGPALFPGGNGVYLMPLQIGADPMAPRM 129
Db 73 -----SNEQNFQLFTHGTIVMLLYGTPIVM-GFANYVLPLOIGADPADVAPRL 119
Qy 130 NNLFSWFLIAGTANGVASLFPAGGQGLGSGVGVLYPPLS--TREAGYSMDLAFIAFVHL 187
Db 120 NAFGMITTVGGVAMLTGFLTPGG----AADFGWTMYSPLSDAIHSPLGSDMWIVGVGA 175
Qy 188 SGASSINGAINMITTFNMRAPGMLTKHKVPLFSWISITAWLIIALLPVLGAIITMLTD 247
Db 176 TGIGSVASANMLTTLICLRAPGTMFRMFIWNIFFVSVLALLIIFPLLLAALGVLYD 235
Qy 248 RNFQTTTFENPAGGDPILYQHILMFFGHPEVYIIILPGFIISHVSTFSKPKVGYLPM 307
Db 236 RKLGGHLYDPANGGS-LLWQHLPFFGHPEVYIILPFGFVSEVPIPVFARKPMFGVYGL 294
Qy 308 VYAMVAICVGLFVVAHMYTVGNSLTQCSYFMLATMVIAPVTGKIFSWIATMGGSV 367
Db 295 IFATLSIGALSMAVVAHMYTVGAVLL--PFESFWTFLISVPTGVKFNWLGTMWRGHIT 352
Qy 368 FKSPMLWAFGMFLFTVGGTVGIVLAQGLDRAYHDTYVVAHMYVNSLGAIFAIFAGI 427
Db 353 WETPMINSVGMATFLEGLTGIMLASPPDLPHLADSYFLIAHFHYTLFGTVVPASCAG 412
Qy 428 YFYNPKFSGRAPPENAAKLHFWTPFPGANVTFFPOHFLGRCMPRRYIDY--PEAFALWN 485
Db 413 YFWEFKMTGRMMDRLKGIHFWLTFVGHGTFELIQHWVGNMGPRIADYLSQGTIYN 472
Qy 486 KVSYSYGAFAPASFLFIVIFYTLVAGERETRPNMGFEADTLEWTLPSPPPAHTFETL 545
Db 473 QISTVFSFLGLSVIPFVNNVFKSVRYGIEIVTVDDPWG-YGNSLEWATSCPPPRHNFASL 531
Qy 546 PK---RSDMDKHPSH 557
Db 532 PRISRPAPELHYPH 547

Search completed: October 18, 2004, 23:38:25
Job time : 134 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 18, 2004, 23:36:01 ; Search time 42 Seconds
(without alignments)
1276.018 Million cell updates/sec

Title: US-09-712-768A-2

Perfect score: 3000

Sequence: 1 MADAAIGHDHHKQGFFTR.....PAHTFTLPKRSDWDKHPSH 557

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: Pir1.*
- 2: Pir2.*
- 3: Pir3.*
- 4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2538.5	84.6	558	2 S08270	cytochrome-c oxida
2	2453	81.8	565	2 S25534	cytochrome-c oxida
3	2347	78.2	554	2 S03809	cytochrome-c oxida
4	1897.5	63.2	572	2 AC2671	cytochrome-c oxida
5	1890.5	63.0	552	2 AC3435	cytochrome-c oxida
6	1873.5	62.5	538	2 S36424	cytochrome-c oxida
7	1847.5	61.6	552	2 D87671	cytochrome-c oxida
8	1829	61.0	541	1 OD231	cytochrome-c oxida
9	1773	59.1	523	2 S62706	cytochrome-c oxida
10	1756.5	58.6	522	2 S25956	cytochrome-c oxida
11	1756.5	58.6	532	2 A97769	cytochrome-c oxida
12	1748.5	58.3	524	2 S14138	cytochrome-c oxida
13	1745	58.2	515	2 T11913	cytochrome-c oxida
14	1744.5	58.1	537	2 S65346	cytochrome-c oxida
15	1744	58.1	534	2 D71698	cytochrome-c oxida
16	1738.5	58.0	527	2 S14139	cytochrome-c oxida
17	1730.5	57.7	524	1 ODR21	cytochrome-c oxida
18	1730.5	57.7	528	1 ODZM1	cytochrome-c oxida
19	1729.5	57.6	524	2 S16256	cytochrome-c oxida
20	1713.5	57.1	514	2 S30291	cytochrome-c oxida
21	1707	56.9	537	2 T11237	cytochrome-c oxida
22	1698.5	56.6	531	2 S78190	cytochrome-c oxida
23	1677.5	55.9	532	2 S59087	cytochrome-c oxida
24	1677	55.9	536	2 S63651	cytochrome-c oxida
25	1670	55.7	533	2 A58931	cytochrome-c oxida
26	1651.5	55.0	527	1 OBSY1	cytochrome-c oxida
27	1643.5	54.8	527	2 S05290	cytochrome-c oxida
28	1642	54.7	531	2 T12406	cytochrome-c oxida
29	1639	54.6	530	2 T11884	cytochrome-c oxida

30 1609 53.6 557 1 ODNCL1
31 1608.5 53.6 527 1 ODOLM
32 1605.5 53.5 567 1 ODAS1
33 1600.5 53.3 471 2 T43827
34 1600 53.3 528 2 S26948
35 1594 53.1 526 2 S62763
36 1594 53.1 541 2 A48327
37 1588.5 52.9 517 2 C34284
38 1564.5 52.1 517 2 S01501
39 1562.5 52.1 517 2 S14205
40 1561.5 52.0 873 1 S53828
41 1561 52.0 519 2 T11129
42 1558.5 51.9 517 2 S70597
43 1558 51.9 516 2 T11182
44 1554.5 51.8 512 1 ODBY1
45 1552 51.7 516 2 T11195

ALIGNMENTS

RESULT 1

S08270
cytochrome-c oxidase (EC 1.9.3.1) chain I-beta - Paracoccus denitrificans
C;Species: Paracoccus denitrificans
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C;Accession: S08270; C35121
R;Raitio, M.; Pispas, J.M.; Metso, T.; Saraste, M.
FEBS Lett. 261, 431-435, 1990
A;Title: Are there isoenzymes of cytochrome c oxidase in Paracoccus denitrificans?
A;Reference number: S08269; MUID:90184495; PMID:2155830
A;Accession: S08270
A;Molecule type: DNA
A;Residues: 1-558 <RAI>
A;Cross-references: UNIPROT:P98002; EMBL:Y07533; NID:G45477; PIDN:CAA68821.1; PID:G45479
R;Van Spanning, R.J.M.; Wansell, C.; Harms, N.; Oltmann, L.F.; Stouthamer, A.H.
J. Bacteriol. 172, 986-996, 1990
A;Title: Mutagenesis of the gene encoding cytochrome c-550 of Paracoccus denitrificans ar
A;Reference number: A35121; MUID:90130336; PMID:2153663
A;Accession: C35121
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-58 <VAN>
A;Cross-references: GB:M27304; NID:G150573; PIDN:AAA8365.1; PID:G150575
R;Buse, G.; Soullamane, T.; Dewor, M.; Meyer, H.E.; Blueggel, M.
Protein Sci. 8, 985-990, 1999
A;Title: Evidence for a copper-coordinated histidine-tyrosine cross-link in the active si
A;Reference number: A58960; MUID:99268331; PMID:10338009
A;Contents: annotation
A;Note: mass spectrographic and chemical characterization of histidyl-tyrosine cross-link
C;Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
C;Keywords: chromoprotein; copper; electron transfer; heme; iron; magnesium; membrane-as
F;27-492/Domain: cytochrome-c oxidase chain I homology <COL>
F;27-59/Domain: transmembrane #status experimental <TM01>
F;84-121/Domain: transmembrane #status experimental <TM02>
F;130-151/Domain: transmembrane #status experimental <TM03>
F;178-206/Domain: transmembrane #status experimental <TM04>
F;218-251/Domain: transmembrane #status experimental <TM05>
F;263-298/Domain: transmembrane #status experimental <TM06>
F;304-322/Domain: transmembrane #status experimental <TM07>
F;334-362/Domain: transmembrane #status experimental <TM08>
F;370-395/Domain: transmembrane #status experimental <TM09>
F;404-430/Domain: transmembrane #status experimental <TM10>
F;441-468/Domain: transmembrane #status experimental <TM11>
F;483-513/Domain: transmembrane #status experimental <TM12>
F;66-80/Disulfide bonds: #status experimental
F;94-413/Binding site: heme a3 iron (His) (axial ligands) #status experimental
F;276-325,326/Binding site: copper (His) #status experimental
F;276-280/Cross-link: 1'-histidyl-3'-tyrosine (His-Tyr) #status experimental
F;280/Binding site: oxygen (Tyr) #status predicted
F;403/Binding site: magnesium (His) (shared with chain II) #status predicted
F;411/Binding site: heme a3 iron (His) (axial ligand) #status experimental

Query Match 84.6%; Score 2538.5; DB 2; Length 558;
 Best Local Similarity 81.0%; Pred. No. 7.2e-167;
 Matches 452; Conservative 51; Mismatches 54; Indels 1; Gaps 1;

QY 1 MADAAIHGH-DHREKQGFTRFMSTNHKDIGLLYLVAAGVGVFISVLFTVMRLLEMDP 59
 Db 1 MADAAIHGH-DHREKQGFTRFMSTNHKDIGLLYLVAAGVGVFISVLFTVMRLLEMDP 60

QY 60 GVQWMCLEGARLIADASQCTANGHLNWNVTYHGIILMMFFVGIIPALFEGGF 119
 Db 61 GVQWMCLEGARLIADASQCTANGHLNWNVTYHGIILMMFFVGIIPALFEGGF 120

QY 120 GAPDMAFPRNNLSFWLFIAGTAMGVASLAPAGGQGLGSGVGVWLYPPLSTREAGYSMD 179
 Db 121 GAPDMAFPRNNLSYWMYVGVASLAPAGGQGLGSGVGVWLYPPLSTREAGYSMD 180

QY 180 LAIPAVHLSGASSINGAINMITTFLNRPAGMTLHKVPLPSWIFITAMLIILALPVLAG 239
 Db 181 LAIPAVHLSGASSILGAINIITITFLNRPAGMTLHKVPLPSWIFITAMLIILALPVLAG 240

QY 240 AITMLLTDNRFGTTFPNPAGGDPILYQHILWFEGHPEVYIIILPGFGLISHVSTFSKK 299
 Db 241 AITMLLMDNRFGTQFPDPAGGDPILYQHILWFEGHPEVYIIILPGFGLISHVSTFSKK 300

QY 300 PVFGYLPVNYAMVAIGLVFVVAHMYTVGMSLTQOSYFMLATMVIATVPTGIKIPSWIA 359
 Db 301 PVFGYLPVNYAMVAIGLVFVVAHMYTVGMSLTQOSYFMLATMVIATVPTGIKIPSWIA 360

QY 360 TMWGSVERKSPMLWAFGEFLETVGVTCIVLAQAGLDRAHYDTHYVVAHMYVMSLGA 419
 Db 361 TMWGSVERKSPMLWAFGEFLETVGVTCIVLAQAGLDRAHYDTHYVVAHMYVMSLGA 420

QY 420 IPAFAGIYFMPKFSGRAPPEAAKLHFWTFPIGANTVFFPQHFLGRCMPRRIDYPE 479
 Db 421 IPAFAGIYFMPKFSGRAPPEAAKLHFWTFPIGANTVFFPQHFLGRCMPRRIDYPE 480

QY 480 AFALNKNVSVYGAFLAFASLFFIYFVTLVAGRETRPNWGPBAPADLEWTLSPPPA 539
 Db 481 AFALNKNVSVYGAFLAFASLFFIYFVTLVAGRETRPNWGPBAPADLEWTLSPPPA 540

QY 540 HTFETLPKRSDDWKHPSH 557
 Db 541 HTFETLPKRSDDWKHPSH 558

RESULT 2

S20534
 cytochrome-c oxidase (EC 1.9.3.1) chain I - Rhodobacter sphaeroides
 C;Species: Rhodobacter sphaeroides
 C;Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 20-Aug-1999
 C;Accession: S20534; A42378
 R;Shapleigh, J.P.; Gennis, R.B.
 Mol. Microbiol. 6, 635-642, 1992

A;Title: Cloning, sequencing and deletion from the chromosome of the gene encoding subunit I of cytochrome c oxidase
 A;Reference number: S20534; PMID:92204019; PMID:1313140
 A;Accession: S20534
 A;Molecule type: DNA
 A;Residues: 1-565 <SH2>
 A;Cross-references: EMBL:X62645; NID:G46417; PIDN:CAA44514.1; PID:G46418
 A;Note: The authors translated the codon AAC for residue 214 as Met and ATG for residue 317, 'AARLN', 'S', '486', 'S', '488', 'S' <SH2>
 R;Shapleigh, J.P.; Hill, J.J.; Alben, J.O.; Gennis, R.B.
 J. Bacteriol. 174, 2338-2343, 1992
 A;Title: Spectroscopic and genetic evidence for two heme-Cu-containing oxidases in Rhodospirillum rubrum
 A;Reference number: A42378; PMID:92202164; PMID:1313003
 A;Accession: A42378
 A;Status: Preliminary; nucleic acid sequence not shown; not compared with conceptual translation
 A;Molecule type: DNA
 A;Residues: 286-289, 'LI', 292-293, 'G', 295-296, 'MI', 299, 'QIVS', 304-305, 'SR', 308-309, 'V', 313-314, 'AARLN', 'S', '486', 'S', '488', 'S' <SH2>
 C;Genetics: ctad
 A;Gene: ctad
 C;Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
 C;Keywords: chromoprotein; copper; electron transfer; heme; iron; magnesium; membrane-as

F;25-500/Domain: cytochrome-c oxidase chain I homology <COI>
 F;102,421/Binding site: heme a iron (His) (axial ligands) #status predicted
 F;284,333/Binding site: copper (His) #status predicted
 F;284,333/Cross-link: 1'-histidyl-3'-tyrosine (His-Tyr) #status predicted
 F;288/Binding site: oxygen (Tyr) #status predicted
 F;411/Binding site: magnesium (His) (shared with chain II) #status predicted
 F;419/Binding site: heme a3 iron (His) (axial ligand) #status predicted

Query Match 81.8%; Score 2453; DB 2; Length 565;
 Best Local Similarity 78.7%; Pred. No. 5.3e-161;
 Matches 446; Conservative 50; Mismatches 59; Indels 12; Gaps 3;

QY 1 MADAAIHGH-DHREKQGFTRFMSTNHKDIGLLYLVAAGVGVFISVLFTVMRLLEMDP 60
 Db 1 MADAAIHGH-DHREKQGFTRFMSTNHKDIGLLYLVAAGVGVFISVLFTVMRLLEMDP 59

QY 61 VQVMCEGARLIADASQCTANGHLNWNVTYHGIILMMFFVGIIPALFEGGF 110
 Db 60 VQVMCEGARLIADASQCTANGHLNWNVTYHGIILMMFFVGIIPALFEGGF 119

QY 111 GNYLMPLOIGAPDMAFPRNNLSFWLFIAGTAMGVASLAPAGGQGLGSGVGVWLYPPLS 170
 Db 120 GNYLMPLOIGAPDMAFPRNNLSFWLFIAGTAMGVASLAPAGGQGLGSGVGVWLYPPLS 179

QY 171 TREAGYSMDLAIPAVHLSGASSINGAINMITTFLNRPAGMTLHKVPLPSWIFITAMLI 230
 Db 180 TREAGYSMDLAIPAVHLSGASSINGAINMITTFLNRPAGMTLHKVPLPSWIFITAMLI 239

QY 231 LLALPVLAGITMLLTDNRFGTTFPNPAGGDPILYQHILWFEGHPEVYIIILPGFGLIS 290
 Db 240 LLALPVLAGITMLLTDNRFGTTFPNPAGGDPILYQHILWFEGHPEVYIIILPGFGLIS 299

QY 291 HVSTFESKPVGVYLPVNYAMVAIGLVFVVAHMYTVGMSLTQOSYFMLATMVIATVPT 350
 Db 300 HVSTFESKPVGVYLPVNYAMVAIGLVFVVAHMYTVGMSLTQOSYFMLATMVIATVPT 359

QY 351 GIKIFSWIATMWGSVERKSPMLWAFGEFLETVGVTCIVLAQAGLDRAHYDTHYVVAH 410
 Db 360 GIKIFSWIATMWGSVERKSPMLWAFGEFLETVGVTCIVLAQAGLDRAHYDTHYVVAH 419

QY 411 FHYVMSLGAFLAFASLFFIYFVTLVAGRETRPNWGPBAPADLEWTLSPPPA 470
 Db 420 FHYVMSLGAFLAFASLFFIYFVTLVAGRETRPNWGPBAPADLEWTLSPPPA 479

QY 471 PRRYIDYPEAFALNKNVSVYGAFLAFASLFFIYFVTLVAGRETRPNWGPBAPADLE 530
 Db 480 PRRYIDYPEAFALNKNVSVYGAFLAFASLFFIYFVTLVAGRETRPNWGPBAPADLE 538

QY 531 WTLSPPPAHTFETLPKRSDDWKHPSH 557
 Db 539 WTLSPPPAHTFETLPKRSDDWKHPSH 565

RESULT 3

S03809
 cytochrome-c oxidase (EC 1.9.3.1) chain I - Paracoccus denitrificans
 C;Species: Paracoccus denitrificans
 C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004
 C;Accession: S03809
 R;Raitio, M.; Jalli, T.; Saraste, M.
 EMBO J. 6, 2825-2833, 1987

A;Title: Isolation and analysis of the genes for cytochrome c oxidase in Paracoccus denitrificans
 A;Reference number: S03809
 A;Accession: S03809
 A;Molecule type: DNA
 A;Residues: 1-554 <RAI>
 A;Cross-references: UNIPROT:P08305; EMBL:X05829; NID:G994806; PIDN:CAA29274.1; PID:G9956;
 C;Genetics:
 A;Gene: COI
 C;Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
 C;Keywords: chromoprotein; copper; electron transfer; heme; iron; magnesium; membrane-as
 F;24-487/Domain: cytochrome-c oxidase chain I homology <COI>
 F;91,408/Binding site: heme a iron (His) (axial ligands) #status predicted

F:273,322,323/Binding site: copper (His) #status predicted
F:273-277/Cross-link: 1'-histidyl-3'-tyrosine (His-Tyr) #status predicted
F:277/Binding site: oxygen (Tyr) #status predicted
F:398/Binding site: magnesium (His) (shared with chain II) #status predicted
F:406/Binding site: heme a3 iron (His) (axial ligand) #status predicted

Query Match 78.2%; Score 2347; DB 2; Length 554;
Best Local Similarity 76.1%; Pred. No. 9.8e-154;
Matches 415; Conservative 59; Mismatches 69; Indels 2; Gaps 1;

Qy 13 EKQGFTRFWSTNKHKIDGLLYLVAAGVGVFISVLFVYVRLMDPGVQVCMLEGARLI 72
Db 11 EKRGFFTFWSTNKHKIDGLYLYLFTAGLAGLISVTLVYVRLMDPGVQVCMLEGARLV 70

Qy 73 ADASOTCTANGHLNVMTYHGLMMPFVGHIPALFGGFGNYLMPLOQTGADNMAPRNNL 132
Db 71 ADAAAECPNLAHLNVNVVYTHGILMMPFVGHIPALFGGFGNYLMPLOQTGADNMAPRNNL 130

Qy 133 SFWLFIAGTANGVASLFAFGDGLGSGVGMVLPPLSTREAGYSMDLAIFAVHLSGASS 192
Db 131 SYWLVCVGSIAIASLLSPGSDQPGAGVGNVLPPLSTREAGYAMDIAIFAVHVSATS 190

Qy 193 IMGAINMTTFLNMRAPGMTLHKVPLFSWTFITAWLILALPVLAGAITMLLTDNFGT 252
Db 191 ILGAINMTTFLNMRAPGMTLHKVPLFAVAVFITAAMILLSLPVLGGITMLLMDNFGT 250

Qy 253 TFEHPAGGDDPILYQHILWFFGHPEVYIILPGFGIISHVSTTSKPVFGYLPVWYAMV 312
Db 251 QFFDPAGGDDPVLVYQHILWFFGHPEVYIILPGFGIISHVSTTSKPVFGYLPVWYAMA 310

Qy 313 AIGVLGFVVAHMYTVGMSLTQSQSYFMLATMVAIVPTGKIFSWIATMCGSVFSPM 372
Db 311 AIAPLGFVVAHMYTVGMSLTQSQSYFQWATMTVAIVPTGKIFSWIATMCGSVFSPM 370

Qy 373 LWAFGFMELFTVGGVTGIVLAQAGLDRAVHDTYVVAHFMVMSLGAIFAFAGIYFMP 432
Db 371 LWALA--FLFTVGGVTGVVIAQGLSDRVYHDTYVVAHFMVMSLGAIFAFAGIYVWG 428

Qy 433 KPSGRAFFEWAAKLFHFTFFIGANTVFPQFLGRQGMRRYIDYPEAFALNKNVSSYGA 492
Db 429 KMSGRQYFEWAQGLHFMFMFGSNLIFPQFLGRQGMRRYIDYPEVFSYWNSSIGA 488

Qy 493 FLAFASLFFVIVFYVTLVAGRRTRPNPWEFADTLWNTLPSPPAHFTTLPKRSDWD 552
Db 489 YISPASLFFVIGIVFTLFAGRVNVNPNVYNEHADTLWNTLPSPPAHFTTLPKRSDWD 548

Qy 553 KHPSH 557
Db 549 RAQAH 553

RESULT 4
AB2671
Cytochrome-c oxidase chain I cox4 [imported] - Agrobacterium tumefaciens (strain C58, Du
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AB2671
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, I.
; Karp, P.; Romero, P.; Grant, C.; Guenther, D.; Kutyavina, T.; Levy, R.; Li, M.; McClell
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
sner, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AB2671
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-572 <KUR>
A:Cross-references: UNIPROT:Q0UHH4; GB:AE008688; PIDN:AAL41784.1; PID:gl7739138; GSPDB:C
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: cox4

A:Map position: circular chromosome
C:Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
C:Keywords: copper; electron transfer; heme; iron; magnesium; membrane-associated complex
F:299,348,349/Binding site: copper (His) #status predicted
F:299-303/Cross-link: 1'-histidyl-3'-tyrosine (His-Tyr) #status predicted
F:303/Binding site: oxygen (Tyr) #status predicted
F:426/Binding site: magnesium (His) (shared with chain II) #status predicted
F:434/Binding site: heme a3 iron (His) (axial ligand) #status predicted

Query Match 63.2%; Score 1897.5; DB 2; Length 572;
Best Local Similarity 63.9%; Pred. No. 7e-123;
Matches 350; Conservative 73; Mismatches 112; Indels 13; Gaps 6;

Qy 7 HGHDRHEKGGFFTRFWSTNKHKIDGLLYLVAAGVGVFISVLFVYVRLMDPGVQVCMCL 66
Db 29 HSDHSHKCGFARWELSTNKHKIDGLYLYLFAIMAGILGGSLSVNRMELQEPGQI--LF 86

Qy 67 EG-ARLIADASOTCTANG--HLNVNMTYHGLMMPFVGHIPALFGGFGNYLMPLOQTGAPD 123
Db 87 HGLASVYGFEGDAADGGKHNFVTTAHLTIMIFFMVPNPMAMIGGFANMIPIMIGAPD 146

Qy 124 MAPRMNLSFWLFIAGTANGVASLFAFGDGLGSGVGMVLPPLSTR-EAGYSMDLAI 182
Db 147 MAPRNLNLSFWLFIAGTANGVASLFAFGDGLGSGVGMVLPPLSTRSGMPGPAVDLAI 206

Qy 183 FAVHLSGASSIMGAINMTTFLNMRAPGMTLHKVPLFSWTFITAWLILALPVLAGAIT 242
Db 207 FSLHVAGASSILGAINMTTFLNMRAPGMTLHKVPLFSWTFITAWLILALPVLAGAIT 266

Qy 243 MLITDRNFGTTFNPPAGGDDPILYQHILWFFGHPEVYIILPGFGIISHVSTTSKPVF 302
Db 267 MLITDRNFGTTFNPPAGGDDPILYQHILWFFGHPEVYIILPGFGIISHVSTTSKPVF 336

Qy 303 GYLPVWYAMVAIGVLFVVAHMYTVGMSLTQSQSYFMLATMVAIVPTGKIFSWIATMW 362
Db 327 GYLGMAYAMVAICAVGFIWAHMYTVGMSLTQSQSYFMLATMVAIVPTGKIFSWIATMW 386

Qy 363 GGSVEFKSPMLAFGMFLFTVGGVTGIVLAQAGLDRAVHDTYVVAHFMVMSLGAIFA 422
Db 387 GGSULTSTPMVAIGRIFLFTVGGVTGIVLAQAGLDRAVHDTYVVAHFMVMSLGAIFA 446

Qy 423 IFAGIYFMPKFSGRAPPEWAAKLFHFTFFIGANTVFPQFLGRQGMRRYIDYPEAF 482
Db 447 IFAGIYFMPKFSGRAPPEWAAKLFHFTFFIGANTVFPQFLGRQGMRRYIDYPEAF 506

Qy 483 LMKVSSYGAFL--AFASLFFVIVFYVTLVAGRRTRPNPWEFADTLWNTLPSPPA 539
Db 507 GWNVSSYGSYISAVAVGIELF---GWAEFAKKRIAGNNPMGEGATTLEMLSSPPY 562

Qy 540 HTPETLPK 547
Db 563 HOWEQLPR 570

RESULT 5
AC3435
Cytochrome-c oxidase (EC 1.9.3.1) [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AC3435
R:DelVecchio, V.G.; Kapral, R.J.; Redkar, R.J.; Patra, G.; Mujer, C.; Ivanova, I.
; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AC3435
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-552 <KUR>
A:Cross-references: UNIPROT:Q8YFQ5; GB:AE008917; PIDN:AAL52646.1; PID:gl7983469; GSPDB:C
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI1465

A:Map position: I

C:Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
 C:Keywords: copper; electron transfer; heme; iron; magnesium; membrane-associated complex
 F:279-328/Binding site: copper (His) #status predicted
 F:279-283/Cross-link: 1'-histidyl-3'-tyrosine (His-Tyr) #status predicted
 F:283/Binding site: oxygen (Tyr) #status predicted
 F:406/Binding site: magnesium (His) (shared with chain II) #status predicted
 F:414/Binding site: heme a3 iron (His) (axial ligand) #status predicted

Query Match 63.0%; Score 1890.5; DB 2; Length 552;
 Best Local Similarity 62.8%; Pred. No. 2e-122;
 Matches 350; Conservative 77; Mismatches 113; Indels 17; Gaps 8;

QY 1 MADAAIH---CHDHEKQGFTHFWSTNKHDKGLLYLVAAGVGVISLVFTVMRLML 57
 DB 1 MAGTAAHEGAHDDHKPHG-WWRVWYSTNKHDKGLTYLFAITAGTIGGALSIAEAEIQ 59
 QY 58 DPGVQVMCLLEG-ARLI--ADASQCTANGHLNMVVTYHGLMMFFVGPALFGGFGNYL 114
 DB 60 EPGIQ--IFHGLAQVYGVGDAALDAGKMFNVFSAHALVMIFFWMPALIGGFANWK 117
 QY 115 MPIOIGAPDWAFFRMNLSFWLFIAGTAMGVASLFAFGGQGLGSGVGVVLYPPLSTR-E 173
 DB 118 VPLMIGAPDWAFFRMNLSFWLFIAGTAMGVASLFAFGGQGLGSGVGVVLYPPLSTR-E 173
 QY 174 AGYSMDLAIFAVHLSGASSIMGAINMTTFLNMRAPGMLHKVLPFSKIFITAMILLIA 233
 DB 178 PGAVDPALIAHLSGASSILGAINFTITLNRAPGMLHKVLPFAVAVLTAFILLUS 237
 QY 234 LPVLGATMLLTDNRFGTTFNPPAGGDPILYQHILWFFGHEPEVYIIILPGFGIISHVV 293
 DB 238 LPVLGATMLLTDNRFGTTFNPPAGGDPILYQHILWFFGHEPEVYIIILPGFGIISHVV 297
 QY 294 STSKRPVFCYLPVYAMVAIGVLGFWWAHMYTVQMSLTQOSYFMTATNVIAPTGK 353
 DB 298 STSKRPVFCYLPVYAMVAIGVLGFWWAHMYTVQMSLTQOSYFMTATNVIAPTGK 357
 QY 354 IFSWIATMGGSVFEXSPMLWAFGMFLFTVGVTGIVLAQGLDRAYHDTYVVAHFY 413
 DB 358 IFSWIATMGGSVFEXSPMLWAFGMFLFTVGVTGIVLAQGLDRAYHDTYVVAHFY 417
 QY 414 VMSLGAIFALFAGIYFMPKSGRAPEWAALKHFWTFPIGANVTFPOHFLGROGMPRR 473
 DB 418 VLSGAVFAIFAGIYFMPKSGRAPEWAALKHFWTFPIGANVTFPOHFLGROGMPRR 477
 QY 474 YIDYPAFAFWNVSSYGAFLAFASLFFIVFVYTL---VAGRRTRNPMGEFADTLE 530
 DB 478 YIDYPAFAFWNVSSYGAFLAFASLFFIVFVYTL---VAGRRTRNPMGEFADTLE 533
 QY 531 WTLPSPPPAHTFTLPK 547
 DB 534 WQLSSPPPPHWEQLPR 550

RESULT 6

S36424
 Cytochrome-c oxidase (EC 1.9.3.1) chain I - Rhizobium leguminosarum
 C:Species: Rhizobium leguminosarum
 C:Date: 09-Dec-1993 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
 C:Accession: S36424
 R:Gabel, C.; Bittinger, M.A.; Maier, R.J.
 submitted to the EMBL Data Library, July 1993
 A:Description: Cytochrome aa3 gene regulation in the Rhizobiaceae: comparison of copper
 A:Reference number: S36424
 A:Accession: S36424
 A:Status: preliminary
 A:Molecule type: DNA
 A:Cross-references: UNIPROT:Q08855; EMBL:X74341; NID:g396698; PID:g3966
 C:Genetics:
 A:Gene: coxA
 C:Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
 C:Keywords: chromoprotein; copper; electron transfer; heme; iron; magnesium; membrane-as

rotein

F:31-481/Domain: cytochrome-c oxidase chain I homology <COI>
 F:82-402/Binding site: heme a iron (His) (axial ligand) #status predicted
 F:265-314,315/Binding site: copper (His) #status predicted
 F:265-269/Cross-link: 1'-histidyl-3'-tyrosine (His-Tyr) #status predicted
 F:269/Binding site: oxygen (Tyr) #status predicted
 F:392/Binding site: magnesium (His) (shared with chain II) #status predicted
 F:400/Binding site: heme a3 iron (His) (axial ligand) #status predicted

Query Match 62.5%; Score 1873.5; DB 2; Length 538;
 Best Local Similarity 64.2%; Pred. No. 2.9e-121;
 Matches 348; Conservative 63; Mismatches 110; Indels 21; Gaps 4;

QY 7 HGHDEHKQGFTHFWSTNKHDKGLLYLVAAGVGVISLVFTVMRLMLMDPGVQMCL 66
 DB 15 HGHDEHAHP---TCWRRSTNKHDKGLTYLFAITAGTIGGALSIAEAEIYMPGVEYF-- 69
 QY 67 EGARLIADASQCTANGHLNMVVTYHGLMMFFVGPALFGGFGNYLMLQIGAPDMAF 126
 DB 70 -----HNTHLNVFVTSHGVIIMFFWMPAMIGGFGNFWFLPLMIGAPDMAF 115
 QY 127 PRMNLSFWLFIAGTAMGVASLFAFGGQGLGSGVGVVLYPPLSTR-EAGYSMDLAIFAV 185
 DB 116 PRMNLSFWLFIAGTAMGVASLFAFGGQGLGSGVGVVLYPPLSTR-EAGYSMDLAIFAV 175
 QY 186 HLCGASSIMGAINMTTFLNMRAPGMLHKVLPFSKIFITAMILLIALPVLGATMLL 245
 DB 176 HLCGASSIMGAINMTTFLNMRAPGMLHKVLPFSKIFITAMILLIALPVLGATMLL 235
 QY 246 TDRNFGTTFNPPAGGDPILYQHILWFFGHEPEVYIIILPGFGIISHVSTFSKRPVGYL 305
 DB 236 TDRNFGTTFNPPAGGDPILYQHILWFFGHEPEVYIIILPGFGIISHVSTFSKRPVGYL 295
 QY 306 PMYAMVAIGVLGFWWAHMYTVQMSLTQOSYFMTATNVIAPTGKIFSWIATMGGGS 365
 DB 296 GMVYAMAAIGLGVVWAHMYTVQMSLTQOSYFMTATNVIAPTGKIFSWIATMGGGS 355
 QY 366 VEFKSPMLWAFGMFLFTVGVTGIVLAQGLDRAYHDTYVVAHFYVMSLGAIFAIPA 425
 DB 356 IEFATPMLWALAFILFTVGVTGIVLAQGLDRAYHDTYVVAHFYVMSLGAIFAIPA 415
 QY 426 GIYPMKSGRAPEPEWAALKHFWTFPIGANVTFPOHFLGROGMPRRYIDYPAFAFWNV 485
 DB 416 GWYTFKSGYMYNETLAEAHFWLFIQVNLIFPFEPHFLGSGMPRRYIDYPAFAFWNV 475
 QY 486 KVSSYGAFLAFASLFFIVFVYTLVAGRRTRNPMGEFADTLEWTLPSPPPAHTFTLP 545
 DB 476 LVSSIGSVISGFSVLLFIYC-VYDAFAKNVGVGNPMWAGATTLWTLPSPPVHFEVL 534
 QY 546 PK 547
 DB 535 PR 536

RESULT 7

DB7671
 Cytochrome-c oxidase (EC 1.9.3.1), subunit I [imported] - Caulobacter crescentus
 C:Species: Caulobacter crescentus
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
 C:Accession: DB7671
 R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.I.
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A:Title: Complete Genome Sequence of Caulobacter crescentus.
 A:Reference number: A87249; MUID:21173698; PMID:11259647
 A:Accession: DB7671
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-552 <STO>
 A:Cross-references: UNIPROT:Q9A300; GB:AE005673; NID:g13425118; PID:AAK35368.1; GSPDB:GN
 C:Genetics:
 A:Gene: CC3406

RESULT 8
ODZJ1
cytochrome-c oxidase (EC 1.9.3.1) chain I - Bradyrhizobium japonicum
C:Species: Bradyrhizobium japonicum
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 09-Jul-2004
C:Accession: S13076; S12101
R:Bott, M.; Bolliger, M.; Hennecke, H.
Mol. Microbiol. 4, 2147-2157, 1990
A:Title: Genetic analysis of the cytochrome c-aa(3) branch of the Bradyrhizobium japonicum
A:Reference number: S13076; MUID:91211625; PMID:1965217
A:Accession: S13076
A:Molecule type: DNA
A:Residues: 1-541 <BOT>
A:Cross-references: UNIPROT:P31833; EMBL:X54800; NID:G39502; PIDN:CAA308570.1; PID:G39503
R:Gabel, C.; Maier, R.J.
Nucleic Acids Res. 18, 6143, 1990
A:Title: Nucleotide sequence of the coxA gene encoding subunit I of cytochrome aa(3) of
A:Reference number: S12101; MUID:91045095; PMID:2172930
A:Accession: S12101

A:Molecule type: DNA
A:Residues: 1-541 <AB>
A:Cross-references: EMBL:X54318; NID:G39505; PIDN:CAA38216.1; PID:G39506
C:Gene: coxa
C:Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
C:Keywords: chromoprotein; copper; electron transfer; heme; iron; magnesium; membrane-asso-
rotein
F:34-484/Domain: cytochrome-c oxidase chain I homology <CO1>
F:41-60/Domain: transmembrane #status predicted <TM01>
F:87-107/Domain: transmembrane #status predicted <TM02>
F:124-142/Domain: transmembrane #status predicted <TM03>
F:173-193/Domain: transmembrane #status predicted <TM04>
F:211-231/Domain: transmembrane #status predicted <TM05>
F:262-282/Domain: transmembrane #status predicted <TM06>
F:294-315/Domain: transmembrane #status predicted <TM07>
F:331-352/Domain: transmembrane #status predicted <TM08>
F:365-385/Domain: transmembrane #status predicted <TM09>
F:404-424/Domain: transmembrane #status predicted <TM10>
F:441-460/Domain: transmembrane #status predicted <TM11>
F:483-503/Domain: transmembrane #status predicted <TM12>
F:85-405/Binding site: heme a iron (His) (axial ligands) #status predicted
F:268 317 318/Binding site: copper (His) #status predicted
F:568-272/Cross-link: 1'-hisidyl-3'-tyrosine (His-Tyr) #status predicted
F:272/Binding site: oxygen (Tyr) #status predicted
F:403/Binding site: heme a3 iron (His) (axial ligand) #status predicted

Query Match 61.0%; Score 1829; DB 1; Length 541;

RESULT 9

```

S62706
Cytochrome-c oxidase (EC 1.9.3.1) chain I - green alga (Platymonas subcordiformis) mitoc
C:Species: mitochondrion Platymonas subcordiformis
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S62706
R:Kessler, U.; Zetsche, K.
Plant Mol. Biol. 29, 1081-1086, 1995
A:Title: Physical map and gene organization of the mitochondrial genome from the unicell
A:Reference number: S62702; MUID:96145517; PMID:855450
A:Accession: S62706
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-523 <RES>
A:Cross-references: UNIPROT:Q36520; EMBL:Z47795; NID:q633584; PIDN:CRAB7753.1; PID:q6335
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995
C:Genetics:
A:Gene: coi
A:Genome: mitochondrion
C:Function:
A:Description: the cytochrome-c oxidase complex catalyzes the oxidation of four molecule
ns from the mitochondrial matrix producing two molecules of water and lowering the conce
A:Pathway: oxidative phosphorylation; respiratory chain
A:Note: chain I directly reduces oxygen on the mitochondrial matrix side of the inner-me
C:Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
C:Keywords: cytochrome; copper; electron transfer; heme; iron; lipoprotein; magnesium
A:Active phosphorylation; oxidoreductase; respiratory chain; transmembrane protein
F:1-13/Domain: mitochondrial matrix #status predicted <MM1>
F:13-460/Domain: cytochrome-c oxidase chain I homology <CO1>
F:13-460/Domain: transmembrane #status predicted <TM01>
F:14-42/Domain: intracristal #status predicted <ITC1>
F:43-54/Domain: transmembrane #status predicted <TM02>
F:55-90/Domain: transmembrane #status predicted <TM02>
F:91-98/Domain: mitochondrial matrix #status predicted <MM2>
F:99-121/Domain: transmembrane #status predicted <TM03>
F:122-144/Domain: intracristal #status predicted <ITC2>
F:145-174/Domain: transmembrane #status predicted <TM04>
F:175-186/Domain: mitochondrial matrix #status predicted <MM3>
F:187-216/Domain: transmembrane #status predicted <TM05>
F:217-231/Domain: intracristal #status predicted <ITC3>
F:232-265/Domain: transmembrane #status predicted <TM06>
F:266-272/Domain: mitochondrial matrix #status predicted <MM4>
F:273-289/Domain: transmembrane #status predicted <TM07>
F:290-301/Domain: intracristal #status predicted <ITC4>
F:302-330/Domain: transmembrane #status predicted <TM08>
F:331-338/Domain: mitochondrial matrix #status predicted <MM5>
F:339-359/Domain: transmembrane #status predicted <TM09>
F:360-373/Domain: intracristal #status predicted <ITC5>
F:374-403/Domain: transmembrane #status predicted <TM10>
F:404-409/Domain: mitochondrial matrix #status predicted <MM6>
F:410-436/Domain: transmembrane #status predicted <TM11>
F:437-449/Domain: intracristal #status predicted <ITC6>
F:450-481/Domain: transmembrane #status predicted <TM12>
F:482-522/Domain: mitochondrial matrix #status predicted <MM7>
F:65,381/Binding site: heme a iron (His) (axial ligands) #status predicted
F:244,293,294/Binding site: copper (His) #status predicted
F:244,248/Cross-link: 1'-histidyl-3'-tyrosine (His-Tyr) #status predicted
F:248/Binding site: oxygen (Tyr) #status predicted
F:322/Binding site: myristate (Lys) (covalent) #status predicted
F:371/Binding site: magnesium (His) (shared with chain II) #status predicted
F:379/Binding site: heme a3 iron (His) (axial ligand) #status predicted

Query Match 59.1%; Score 1773; DB 2; Length 523;
Best Local Similarity 60.1%; Pred. No. 2,2e-114;
Matches 322; Conservative 82; Mismatches 104; Indels 28; Gaps 6;

QY 17 FFFRWFSTNHHKDIGLLYLVAAGVGVFISVLFTVYVNRLEMDPGVYMCLEGARLIADAS 76
DB 4 FAQRLWLFSTNHHKDIGLLYLVAAGVGVFISVLFTVYVNRLEMDPGVYMCLEGARLIADAS 48
QY 77 QTCTANGHLNWNVYTHGILMMEFVGIPALFGFGNGYLMPLQIGAPDMAPPRNNLSFWL 136
DB 49 QVLGNHQLNVLITAHAFIMFFVWMPALJGFGNGVFPIMIGAPDMAPPRNNLSFWL 108
QY 137 FIAGTANGVASLFPAGGGDGLGSGGVNLYPPILST--REAGYSMDLAIFAVHLSGASSIM 194

```

Db 109 LPPSLLLSSALV-----EVGCGSGWTVPPLSGITSHSGGSDVLAIFSLHLSGVSSIL 163
Qy 195 GAINMITTFLNRRAGMTLHKVPLFSWGFITAMILLALPVLGAIYMLLTDNRFGTTF 254
Db 164 GINFTITFNNRPAFLMHRLEPLFVNSVLVAFLLLSLPLVLAGAITMLLTDNRNFTF 223
Qy 255 FNPAGGDPILYQHILWFFGHEPEVYIILDPGFIISHVVSTFSKPKPVGYLPMYAMAI 314
Db 224 FDPAGGDPILYQHILWFFGHEPEVYIILDPGFIISHIVSTFSRKPVPFGYLGMYAMISI 283
Qy 315 GVLGFVVAHHNYTVQMSLTQOSYFMLATWIAVPTGKIISWTATWGGSVFKSPMLW 374
Db 284 GVLGFVVAHHNYTVQMSLTQOSYFMLATWIAVPTGKIISWTATWGGSIQYKTEMLF 343
Qy 375 AFGFMFLFTVGGVGTGIVLAQGLDRAYHDTYVVAHFHYVMSLGAIFAFAFAGIYFMPKF 434
Db 344 AVGFIFLFTVGGVGTGIVLAQGLDRAYHDTYVVAHFHYVMSLGAIFAFAFAGIYFMPKF 403
Qy 435 SGRAPEWAUKHFWTFFIGANVTFFPOHFLGRQGNPRYIDYPAFALMNKVSYSYAGFL 494
Db 404 TGLQYPETLQGHFWITFFGNLTFFPMHFLGLAGMPRIIDYDPDAYAGWNAFSSFGSYV 463
Qy 495 AFAS-FLRFIVFVTLVAGRETRPNWG--EFADTLEWTLPSPPAHTTETLP 546
Db 464 SVVGIFCFVTVVL--TLTSENKCAPSPWAVEQNSTLTLEWVPSPPAHTTETLP 516
RESULT 11
cytochrome-c oxidase (EC 1.9.3.1) - Rickettsia conorii (strain Malieh 7)
C:Species: Rickettsia conorii
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Rickettsia conorii
Science 293, 2093-2098, 2001
A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A:Reference number: A97700; MUID:21442074; PMID:11557893
A:Accession: A97769
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-532 <KUR>
A:Cross-references: UNIPROT:Q92167; GB:AE006914; PIDN:AAL03091.1; PID:G15619633; GSPDB:G15619633
A:Gene: coxI
C:Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
C:Keywords: copper; electron transfer; heme; iron; magnesium; membrane-associated complex
F:79,395/Binding site: heme a iron (His) (axial ligands) #status predicted
F:258,307/Binding site: copper (His) #status predicted
F:258-262/Cross-link: 1'-histidyl-3'-tyrosine (His-Tyr) #status predicted
F:262/Binding site: oxygen (Tyr) #status predicted
F:385/Binding site: magnesium (His) (shared with chain II) #status predicted
F:393/Binding site: heme a3 iron (His) (axial ligand) #status predicted
Query Match 58.6%; Score 1756.5; DB 2; Length 532;
Best Local Similarity 59.3%; Pred. No. 3.1e-113;
Matches 320; Conservative 78; Mismatches 119; Indels 23; Gaps 4;
Qy 9 HDHHEKQGFTEWSTNKHDKGLYLVAAGVGVGFISVLFTVYMLELMDPQVQVQVLEG 68
Db 10 HDGHTPHGWRRLSTNKHDKGLYLVAAGVGVGFISVLFTVYMLELMDPQVQVQVLEG 65
Qy 69 ARLIADASOTCANGHLMNVMYTHGILMMFVGVIPALFGGFGNYLMPLOIGAPDMPAFPR 128
Db 66 -----NHDFOLYNLVITAHAVIMVFWMIPALFGGFGNYLMPLOIGAPDMPAFPR 114
Qy 129 MNLSFWLIAGTAMGVASLFPAGDGGQLGSGVGVWLYPPLS--TREAGYSMDLAIFAIVH 186
Db 115 LNNISFWLIVPAFLLMGSAFVDG-----GPGTGTWLYPPLSLNLSGHPCAAVDMAIFSLH 169
Qy 187 LGCASSIMGAINMITTFLNRRAGMTLHKVPLFSWGFITAMILLALPVLGAIYMLLTDNR 246
Db 170 L'GLSILGSLINLIVTFNNRPAFLMHRLEPLFVNSVLVAFLLLSLPLVLAGAITMLLTDNR 229

Qy 247 DRNFGTTFNPAGGDPILYQHILWFFGHEPEVYIILDPGFIISHVVSTFSKPKPVGYLP 306
Db 230 DRNFGTTFNPAGGDPILYQHILWFFGHEPEVYIILDPGFIISHVVSTFSKPKPVGYLP 289
Qy 307 MYAMVAIGVLGFVVAHHNYTVQMSLTQOSYFMLATWIAVPTGKIISWTATWGGSV 366
Db 290 MIGAMVIIGVGVFVVAHHNYTVQMSLTQOSYFMLATWIAVPTGKIISWTATWGGSI 349
Qy 367 BFKSPMLWAFGPMFLFTVGGVGTGIVLAQGLDRAYHDTYVVAHFHYVMSLGAIFAFAFAG 426
Db 350 TFFTPMLFSGFIIILFTTIGGVTGIIISNSALDRVLDHDTYVVAHFHYTWSLGLFTAFAG 409
Qy 427 IYFMPKESGRAPPEWAUKHFWTFFIGANVTFFPOHFLGRQGNPRYIDYPAFALMNK 486
Db 410 FYWFGKISQKQYDPDILGKTHFWITFFGNLTFFPMHFLGLAGMPRIIDYDPDAYAGW 469
Qy 487 VSSVGAFLAFAFLFTFVIFVYTLVAGERETRNPNWGEFADTLEWTLPSPPAHTTETLP 546
Db 470 VSSIGAGISWFAALYFVIFVYTLKYG-KDCFPNPMWEGADTLEWTLSPPPHTTETTP 528
RESULT 12
SL14138
cytochrome-c oxidase (EC 1.9.3.1) chain I - sugar beet mitochondrion
C:Species: mitochondrion Beta vulgaris var. altissima (sugar beet)
C>Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 09-Jul-2004
C:Accession: SL14138
R:Sendra, M.; Harada, T.; Mikami, T.; Sugitara, M.; Kinoshita, T.
Curr. Genet. 19, 175-181, 1991
A:Title: Genomic organization and sequence analysis of the cytochrome oxidase subunit II
A:Reference number: SL14138; MUID:91330331; PMID:1651175
A:Accession: SL14138
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-524 <SEN>
A:Cross-references: UNIPROT:P24794; GB:X57693; GB:S47702; NID:g11258; PIDN:CAA40874.1; PJ
C:Genetics:
A:Gene: coxII
A:Genome: mitochondrion
C:Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
C:Keywords: chromoprotein; copper; electron transfer; heme; iron; magnesium; membrane-associated protein
F:12-459/Domain: cytochrome-c oxidase chain I homology <COI>
F:64,380/Binding site: heme a iron (His) (axial ligands) #status predicted
F:243,282,293/Binding site: copper (His) #status predicted
F:243-247/Cross-link: 1'-histidyl-3'-tyrosine (His-Tyr) #status predicted
F:247/Binding site: oxygen (Tyr) #status predicted
F:370/Binding site: magnesium (His) (shared with chain II) #status predicted
F:378/Binding site: heme a3 iron (His) (axial ligand) #status predicted
Query Match 58.3%; Score 1748.5; DB 2; Length 524;
Best Local Similarity 59.2%; Pred. No. 1.1e-112;
Matches 315; Conservative 86; Mismatches 106; Indels 25; Gaps 5;
Qy 20 RWFMTNKHDKGLYLVAAGVGVGFISVLFTVYMLELMDPQVQVQVQVLEGADIASQTC 79
Db 6 RLPLSTNKHDKGLYTFYFGAIVGVTGCTFSLRMELARPG-----DQIL 50
Qy 80 TANGHLMNVMYTHGILMMFVGVIPALFGGFGNYLMPLOIGAPDMPAFPRMNNLSFWL 139
Db 51 GGNHQLYNLVITAHAFIMVFWMIPALFGGFGNYLMPLOIGAPDMPAFPRMNNLSFWL 110
Qy 140 GTAMGVASLFPAGDGGQLGSGVGVWLYPPLS--TREAGYSMDLAIFAIVHLSGASSIMGAI 197
Db 111 SLULLLSALV-----EVGSGTGTWLYPPLSGITSHSGGAVDLAIFSLHLSGVSSILGSI 165
Qy 198 NMITFLNRRAGMTLHKVPLFSWGFITAMILLALPVLGAIYMLLTDNRNFTTFFNP 257
Db 166 NFITTFNNRPGMTLHKVPLFSWGFITAMILLALPVLGAIYMLLTDNRNFTTFFNP 225
Qy 258 AGGSDPILYQHILWFFGHEPEVYIILDPGFIISHVVSTFSKPKPVGYLPMYAMVAIGVL 317
Db 226 AGGSDPILYQHILWFFGHEPEVYIILDPGFIISHVVSTFSKPKPVGYLPMYAMVAIGVL 285

```

QY 318 GFVVAAHMYTVGMSLTQOOSYFVLATWIAVPTGKIFSWIATMWSGSEYKSPMLWAPG 377
DB 286 GFLVAAHMYTVGMSLTQOOSYFVLATWIAVPTGKIFSWIATMWSGSEYKSPMLWAPG 345
QY 378 FMFLFTVGGVTGVLAAQAGLDRAHYDHTYYVVAHFHYVMSLGAIFAIFAGIYFVMPKFSR 437
DB 346 FIELFTVGGTGLVLANSGLDIALHDTYYVVAHFHYVMSLGAIFAIFAGIYFVMPKFSR 405
QY 438 AFPEWAAKLFHMTFFIGANVTFFPQHLGRQGMPPRYIDYPEAFALWNKVSYGAFALAPA 497
DB 406 TYPETLQIHFWLMTFFPGVNVNFTFPMFLGSLGMPRIIPDYPAAYAGNALSGSGSVISV 465
QY 498 SFLFFVIFVYITLVAGR-RETENPAG--EFADTLEWTLPSPPPAHTFTLTP 546
DB 466 GICCFVFWITLSSGKNKRCAPSNAVEENSTLEWMLPSPAPHTFGELP 517

RESULT 13
T11913
cytochrome-c oxidase (EC 1.9.3.1) chain I - Prototheca wickerhamii mitochondrion
C/Species: mitochondrion Prototheca wickerhamii
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C/Accession: T11913
R/Wolff, G.; Plante, I.; Lang, B.F.; Kueck, U.; Burger, G.
J. Mol. Biol. 237, 75-86, 1994
A/Title: Complete sequence of the mitochondrial DNA of the chlorophyte alga Prototheca
A/Reference number: Z17373; MUID:94180393; PMID:8133522
A/Accession: T11913
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-515 <MOL>
A/Cross-references: UNIPROT:Q37616; EMBL:U02970; NID:g467843; PID:g467847; PIDN:AAD12634
A/Experimental source: strain 263-11
C/Genetics:
A/Genome: mitochondrion
A/Introns: 129/2; 237/1; 240/3
A/Note: cox1
C/Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
C/Keywords: copper; electron transfer; heme; iron; magnesium; membrane-associated complex
F/10-457/Domain: cytochrome-c oxidase chain I homology <COI>
F/62,378/Binding site: heme a iron (His) (axial ligands) #status predicted
F/241,290,291/Binding site: copper (His) #status predicted
F/241-245/cross-link: 1'-histidyl-3'-tyrosine (His-Tyr) #status predicted
F/245/Binding site: oxygen (Tyr) #status predicted
F/368/Binding site: magnesium (His) (shared with chain II) #status predicted
F/376/Binding site: heme a3 iron (His) (axial ligand) #status predicted

Query Match 58.2%; Score 1745; DB 2; Length 515;
Best Local Similarity 58.2%; Pred. No. 1.8e-112;
Matches 310; Conservative 92; Mismatches 103; Indels 28; Gaps 6;

QY 19 TRFMSTNKHDKGLLILVAAAGVGVTSVLTFTVMRLMLMDPGVQVWCLSGARLIADASQT 78
DB 3 TRWLYSTNKHDKITGMVLIIFAGSGVLTGTFSLIRMELAQPG-----N01 47
QY 79 CTANGHLNWNVYTHGILNMFYFIPALFGGNGYLMPLQIAPDMAPFMNLSFWLPI 138
DB 48 LNHQHLNLYNIIITAHAFMLTFFMLPALMGFGNFWFLPILIGAPDMAPFMNLSFWLPI 107
QY 139 AGTAMGVASLPAFGDQGLSGGVWVLYPPLST--REAGYSMDLAFPAVHLSGASIMGA 196
DB 108 PSLLLLVSSALV-----EVGAGTGWTVPPLASIASHSGSGVDLAFSLHLAGVSSILGA 162
QY 197 INMITPLNMRAPGWLTHKVPFWSWIFITAWLILLALPVLAGAIFMLLTDNRNFGTTFN 256
DB 163 INFICTVFNMRACGMSHRLPLFWNAVITAWLILLCLPLVLAGIITMLTDNRNFTSPFD 222
QY 257 PAQGGDPILYQHILWFFGHPEVVIILPGGFIISHVYTSFKKPVGYLPMYAMVAIGV 316
DB 223 PAQGGDPILYQHILWFFGHPEVVIILPGGFIISHVYTSFKKPVGYLPMYAMVAIGV 282
QY 317 LGFVAAHMYTVGMSLTQOOSYFVLATWIAVPTGKIFSWIATMWSGSEYKSPMLWAP 376

```

```

DB 283 LGFVAAHMYTVGMSLTQOOSYFVLATWIAVPTGKIFSWIATMWSGSEYKSPMLFAV 342
QY 377 FMFLFTVGGVTGVLAAQAGLDRAHYDHTYYVVAHFHYVMSLGAIFAIFAGIYFVMPKFSR 436
DB 343 GFLFTVGGTGLVLANSGLDVAFHDTYYVVAHFHYVMSLGAIFAIFAGIYFVMPKFSR 402
QY 437 RAFFPEWAAKLFHMTFFIGANVTFFPQHLGRQGMPPRYIDYPEAFALWNKVSYGAFALAP 496
DB 403 LOYPETLQIHFWLMTFFPGVNVNFTFPMFLGSLGMPRIIPDYPAAYAGNALSGSGSVLSI 462
QY 497 ASFLFFVIFVYITLVAGR-RETENPAG--EFADTLEWTLPSPPPAHTFTLTP 545
DB 463 TAVLFFVYVYKTLTS--NEVCPNPNWETTPGVSPTLEWMLPSPAPHTFEEI 513

RESULT 14
S65346
cytochrome-c oxidase (EC 1.9.3.1) chain I - tomato mitochondrion
C/Species: mitochondrion Lycopersicon esculentum (tomato)
C/Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C/Accession: S65346; S65080
R/Kadowaki, K.; Ozawa, K.; Kazama, S.; Kubo, N.; Akihama, T.
Curr. Genet. 28, 415-422, 1995
A/Title: Creation of an initiation codon by RNA editing in the coxI transcript from tomato
A/Reference number: S65080; MUID:96155620; PMID:8575013
A/Accession: S65346
A/Molecule type: mRNA
A/Residues: 1-527 <KAD>
A/Cross-references: UNIPROT:Q34882; EMBL:X54738
A/Experimental source: strain Ponderosa; leaf
A/Note: 1-Met, 4-Leu, 85-Phe, 151-Phe, 194-Leu, 197-Leu, 223-Phe, 249-Leu, 254-Phe, 396-J
A/Accession: S65080
A/Molecule type: DNA
A/Residues: 'T', 2-3, 'P', 5-84, 'S', 86-150, 'S', 152-183, 'S', 185-196, 'P', 198-222, 'S', 224-248, '
A/Cross-references: EMBL:X54738; NID:g1165117; PIDN:CAA38535.1; PID:g1165118
A/Experimental source: strain Ponderosa; leaf
C/Genetics:
A/Genes: COXI
A/Genome: mitochondrion
C/Function:
A/Description: the cytochrome-c oxidase complex catalyzes the oxidation of four molecules
ns from the mitochondrial matrix producing two molecules of water and lowering the concen
A/Pathway: oxidative phosphorylation; respiratory chain
A/Note: Chain I directly reduces oxygen on the mitochondrial matrix side of the inner-mem
C/Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
C/Keywords: chromoprotein; copper; electron transfer; heme; iron; lipoprotein; magnesium;
active phosphorylation; oxidoreductase; respiratory chain; RNA editing; transmembrane prot
F/1-12/Domain: mitochondrion matrix #status predicted <MM1>
F/12-459/Domain: cytochrome-c oxidase chain I homology <COI>
F/13-41/Domain: transmembrane #status predicted <TM01>
F/42-53/Domain: intracrystal #status predicted <ITC1>
F/54-89/Domain: transmembrane #status predicted <TM02>
F/90-97/Domain: mitochondrion matrix #status predicted <MM2>
F/98-120/Domain: transmembrane #status predicted <TM03>
F/121-143/Domain: intracrystal #status predicted <ITC2>
F/144-173/Domain: transmembrane #status predicted <TM04>
F/174-185/Domain: mitochondrion matrix #status predicted <MM3>
F/186-215/Domain: transmembrane #status predicted <TM05>
F/216-230/Domain: intracrystal #status predicted <ITC3>
F/231-264/Domain: transmembrane #status predicted <TM06>
F/265-271/Domain: mitochondrion matrix #status predicted <MM4>
F/272-288/Domain: transmembrane #status predicted <TM07>
F/289-300/Domain: intracrystal #status predicted <ITC4>
F/301-329/Domain: transmembrane #status predicted <TM08>
F/330-337/Domain: mitochondrion matrix #status predicted <MM5>
F/338-358/Domain: transmembrane #status predicted <TM09>
F/359-372/Domain: intracrystal #status predicted <ITC5>
F/373-402/Domain: transmembrane #status predicted <TM10>
F/403-408/Domain: mitochondrion matrix #status predicted <MM6>
F/409-435/Domain: transmembrane #status predicted <TM11>
F/436-448/Domain: intracrystal #status predicted <ITC6>
F/449-480/Domain: transmembrane #status predicted <TM12>

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 18, 2004, 23:28:06 ; Search time 197 Seconds
(without alignments)
1626.821 Million cell updates/sec

Title: US-09-712-768A-2

Perfect score: 3000

Sequence: 1 MADAAIHGDHHEKQGFFTR.....PAHIFETLPKRSDDWKHPSH 557

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_02:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB	ID	Description
1	2538.5	84.6	558	1	CX1B_PARDE	P98002 paracoccus
2	2493.5	83.1	566	1	CX1 RHOSH	P33517 rhodobacter
3	2347	78.2	554	1	CX1A_PARDE	P08305 paracoccus
4	1898.5	63.3	562	2	Q92RG9	O92RG9 rhizobium m
5	1897.5	63.2	572	2	Q8YHB4	Q8YHB4 agrobacteri
6	1891.5	63.0	552	2	Q8Q263	Q8Q263 brucella su
7	1890.5	63.0	552	2	Q8TFQS	Q8TFQS brucella me
8	1873.5	62.5	538	1	CX1 RHILE	Q08855 rhizobium l
9	1847.5	61.6	552	2	Q9A300	Q9A300 caulobacter
10	1846.5	61.6	550	2	Q985X1	Q985X1 rhizobium l
11	1842	61.4	541	2	Q6NEJ6	Q6NEJ6 rhodosphe
12	1842	61.4	541	2	CAE26276	CAE26276 rhodosphe
13	1838	61.3	539	2	Q59631	Q59631 nitrobacter
14	1829	61.0	541	1	CX1 BRAJA	P31833 bradyrhizob
15	1773	59.1	523	2	Q36520	Q36520 platymonas
16	1762.5	58.8	538	2	Q8W9R3	Q8W9R3 mesostigma
17	1762.5	58.6	522	1	CX1 MARPO	P26856 marchantia
18	1756.5	58.6	525	2	Q9TCA6	Q9TCA6 nephrolep
19	1756.5	58.6	532	2	Q92167	Q92167 rickettsia
20	1751.5	58.4	525	2	Q63228	Q63228 populus tre
21	1748.5	58.3	524	1	CX1 BETVU	P24794 beta vulgar
22	1748.5	58.3	524	2	Q7HJM3	Q7HJM3 beta vulgar
23	1748.5	58.3	524	2	BAA86630	BAA86630 beta vulg
24	1748.5	58.3	524	2	BAA99455	BAA99455 beta vulg
25	1748.5	58.3	532	2	Q7PBR5	Q7PBR5 rickettsia
26	1746.5	58.2	526	2	Q8M1D5	Q8M1D5 chaetosphe
27	1745	58.2	515	2	Q37616	Q37616 prototheca
28	1744	58.1	534	1	CX1 RICPR	O54069 rickettsia
29	1739.5	58.0	524	2	Q7YAN3	Q7YAN3 chara vulga
30	1738.5	58.0	527	1	CX1 ARATH	P60620 arabidopsis
31	1738.5	58.0	527	1	CX1 RAPSA	P60621 rapanus sa

RESULT 1

ID CX1B_PARDE STANDARD; PRT; 558 AA.
AC P98002;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 01-OCT-2004 (Rel. 45, Last annotation update)

DE Cytochrome c oxidase polypeptide I-beta (EC 1.9.3.1) (Cytochrome aa3

DE subunit I-beta).

GN Name=ctadII;

OS Paracoccus denitrificans.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacteriales;

OC Rhodobacteraceae; Paracoccus.

OX NCBI_TaxID=266;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=PD 1222;

RX MEDLINE=90184495; PubMed=2155830;

RA Raitio M., Pispas J.M., Metso T., Saraste M.;

RT "Are there isoenzymes of cytochrome c oxidase in Paracoccus

denitrificans?";

RL FEBS Lett. 261:431-435(1990).

RN [2]

RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).

RX MEDLINE=95379947; PubMed=7651515;

RA Iwata S., Ostermeier C., Ludwig B., Michel H.;

RT "Structure at 2.8 A resolution of cytochrome c oxidase from Paracoccus

denitrificans.";

RL Nature 376:660-669(1995).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).

RX MEDLINE=98021406; PubMed=9380672;

RA Ostermeier C., Harrenga A., Ermler U., Michel H.;

RT "Structure at 2.7-A resolution of the Paracoccus denitrificans two-

subunit cytochrome c oxidase complexed with an antibody FV fragment.";

Proc. Natl. Acad. Sci. U.S.A. 94:10547-10553(1997).

RN [4]

RP COVALENT BOND.

RX MEDLINE=99268331; PubMed=10338009;

RA Buse G., Soulliane T., Dewor M., Meyer H.E., Blueggel M.;

RT "Evidence for a copper-coordinated histidine-tyrosine cross-link in

the active site of cytochrome oxidase.";

Protein Sci. 8:985-990(1999).

RN [5]

RP REVIEW.

RX MEDLINE=96391941; PubMed=8794157;

RA Ostermeier C., Iwata S., Michel H.;

RT "Cytochrome c oxidase";

RL Curr. Opin. Struct. Biol. 6:460-466(1996).

CC -!- FUNCTION: Subunit I and II form the functional core of the enzyme complex. Electrons originating in cytochrome c are transferred via heme a and Cu(A) to the binuclear center formed by heme a3 and Cu(B). This cytochrome c oxidase shows proton pump activity across the membrane in addition to the electron transfer.

ALIGNMENTS

Q6WAV9 brassica ju
Q6YSM3 brassica na
AAP58355 brassica
BAC98915 brassica
Q7YFEL brassica ju
P14578 cryza sativ
Q7JAI6 cryza sativ
BAC19897 oryza sat
P08742 zea mays (m
AAN33123 zea mays
AAR91047 zea mays
P05502 sorghum bic
Q35812 sorghum bic
P08741 triticum ae

32 1738.5 58.0 527 2 Q6WAV9
33 1738.5 58.0 527 2 Q6YSM3
34 1738.5 58.0 527 2 AAP58355
35 1738.5 58.0 527 2 BAC98915
36 1738.5 57.8 527 2 Q7YFEL
37 1730.5 57.7 524 1 COX1_CRYSA
38 1730.5 57.7 524 2 Q7JAI6
39 1730.5 57.7 524 2 BAC19897
40 1730.5 57.7 528 2 COX1_MAIZE
41 1730.5 57.7 528 2 AAN33123
42 1730.5 57.7 528 2 AAR91047
43 1730.5 57.7 530 1 COX1_SORBI
44 1730.5 57.7 632 2 Q35812
45 1729.5 57.6 524 1 COX1_WHEAT

[illegible]

Db 301 PFGYLPVLAWAIGILGVVVAHHMYTAGMSUTQQAIFVLAATMTAVPTGKIVFWSIA 360
 QV 360 TMWGSVEFKSPMLWAFGMFLFTVGGVTGIVLAQAGLDRAHYDTYVVAHFHYVMSLGA 419
 Db 361 TMWGSIEFKPMLWAFGLFLFTVGGVTGVLSQAPLDLYHYDTYVVAHFHYVMSLGA 420
 QV 420 IPATAGIYFYMFKSGRAFPPEWAALHFWTFFIGANVTFFPQHFGLGRQMPRIYDIPE 479
 Db 421 VFGIPAGYVYWGKSGRQYPEWAGQLHFWMMFIGNLIFFPQHFGLGRQMPRIYDYPV 480
 QV 480 AFALWKNVSSYGAFLAFASFLFFVIFVYTVLVAGRRETRPNPWEADTLEWTLSPPPA 539
 Db 481 EFAYWNNSSGAYISFASFLFFVIFVYTVLVAGRVNVVYVNEHADTLEWTLSPPP 540
 QV 540 HTFETPLPKRSDWKHPSH 557
 Db 541 HTFETPLPKREDWDRAHAH 558

RESULT 2

COX1_RHOSH
 ID COX1_RHOSH STANDARD; PRT; 566 AA.
 AC P33517;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Cytochrome c oxidase polypeptide I (EC 1.9.3.1) (Cytochrome a3 subunit 1).
 GN Name=ctaD;
 OS Rhodobacter sphaeroides (Rhodospseudomonas sphaeroides).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
 OC Rhodobacteraceae; Rhodobacter.
 OX NCBI_taxID=1063;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GA;
 RX MEDLINE=92204019; PubMed=1313140;
 RA Shapleigh J.P., Gennis R.B.;
 RT "Cloning, sequencing and deletion from the chromosome of the gene encoding subunit I of the a3-type cytochrome c oxidase of Rhodobacter sphaeroides";
 RL Mol. Microbiol. 6:635-642(1992).
 RN [2]
 RP REVISIONS TO 436-439 AND 518-521.
 RA Shapleigh J.P., Gennis R.B.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory chain that catalyzes the reduction of oxygen to water. Subunits 1-3 form the functional core of the enzyme complex. Co I is the catalytic subunit of the enzyme. Electrons originating in cytochrome c are transferred via the copper A center of subunit 2 and heme a of subunit 1 to the bimetallic center formed by heme a3 and copper B. This cytochrome c oxidase shows proton pump activity across the membrane in addition to the electron transfer.
 CC -!- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferricytochrome c + 2 H(2)O.
 CC -!- COFACTOR: Two heme groups and copper B.
 CC -!- PATHWAY: Respiratory chain; terminal step.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.

This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; X62645; CAA44514.2; -
 DR PDB; 1M56; X-ray; A/G=1-566.
 DR PDB; 1M57; X-ray; A/G=1-566.
 DR InterPro; IPR000883; COX1.

DR Pfam; PF00115; COX1; 1.
 DR PRINTS; PR01165; CYCOXIDASE1.
 DR PROSITE; PS0855; COX1; 1.
 DR PROSITE; PS00077; COX1_CUB; 1.
 KW 3D-structure; Copper; Heme; Hydrogen ion transport; Oxidoreductase;
 KW Respiratory chain; Transmembrane.
 FT METAL 102 102 Iron (heme A axial ligand) (Probable).
 FT METAL 284 284 Copper B (Probable).
 FT METAL 288 288 Copper B (Probable).
 FT METAL 333 333 Copper B (Probable).
 FT METAL 334 334 Copper B (Probable).
 FT METAL 419 419 Iron (heme A3 axial ligand) (Probable).
 FT METAL 421 421 Iron (heme A axial ligand) (Probable).
 FT METAL 284 288 1'-histidyl-3'-cyrosine (His-Tyr) (By similarity).
 SQ SEQUENCE 566 AA; 63147 MW; 65A74DBCCSC550B0 CRC64;
 Query Match 83.1%; Score 2493.5; DB 1; Length 566;
 Best Local Similarity 79.4%; Pred. No. 3.7e-153;
 Matches 450; Conservative 49; Mismatches 57; Indels 11; Gaps 2;
 QY 1 MADAALHGHHEKQGFTRFMSTNHKDIGLLYLVAAGVGFISVLFTVYVRLELMDPG 60
 Db 1 MADAALHGHHE-HDRRGFFTRFMSTNHKDIGLLYLFTGGLVGLISVAFVYVRMELMAPG 59
 QY 61 VQWMCLEGAR-----LIADASQCTANGHLWNWVTVYHGIILMFFVGPALFGGF 110
 Db 60 VQWCAEHLESGLVKGFOSLWPSAVENCTPNGLWNWVITGHILMFFVVLPAFGGF 119
 QY 111 GNYLMLPQIGAPDMAFPRMNNLSFWLFIAGTAGVASLFPAGDGGQLGSGVGNLYPPLS 170
 Db 120 GNYFMPLHIGAPDMAFPRMNNLSYWLIVAGTSLAVASLFPAGGNGQLGSGIGVLYPPLS 179
 QY 171 TREAGYSMDLAIPAVHLSGASSIMGAINMTTFLNNRAPGMTLHKYPLFSWSIFITAWLI 230
 Db 180 TSESGYSTDLAIPAVHLSGASSILGAINMTTFLNNRAPGMTLHKYPLFAWSIFITAWLI 239
 QY 231 LLALPVLAGAITMLLTDNRNFTTFNPAGGDPILYQHILWFFGHDPVYIIILPFGGIIS 290
 Db 240 LLALPVLAGAITMLLTDNRNFTTFNPAGGDPVLYQHILWFFGHDPVYIIILPFGGIVS 299
 QY 291 HVSTFSKPKPVFGYLPNVYAMVAIGVLGVVVAHHMYTVGMSLTQOSYFMLATMTAVPT 350
 Db 300 HVIATFAKPKPIFGYLPNVYAMVAIGVLGVVVAHHMYTAGLSLTQOSYFMMATMTAVPT 359
 QY 351 GIKIFSWIATMWGSSVEFKSPMLWAFGMFLFTVGGVTGIVLAQAGLDRAHYDTYVVAH 410
 Db 360 GIKIFSWIATMWGSSIELKTPMLWALGFLFTVGGVTGIVLSQASVDYHYDTYVVAH 419
 QY 411 FHYVMSLGAIFAIFAGIYFYMPKESGRAPPEWAALHFWTFFIGANVTFFPQHFGLGRQM 470
 Db 420 FHYVMSLGAIVFGIFAGIYFWIGKSGRQYPEWAGQLHFWMMFVGNLTFPQHFGLGRQM 479
 QY 471 PRYIDYPEAFALWKNVSSYGAFLAFASFLFFVIFVYTVLVAGRRETRPNPWEADTLE 530
 Db 480 PRYIDYPEAFAFWNNFVSSYGAFLAFASFLFFVIFVYTVLVAGRVNVVYVNEHADTLE 539
 QY 531 WTLSPPPPAHTPETLPKRSDWKHPSH 557
 Db 540 WTLSPPPPEHTPEQLPKREDWERAPAH 566

RESULT 3

CX1A_PARDE
 ID CX1A_PARDE STANDARD; PRT; 554 AA.
 AC P08305;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Cytochrome c oxidase polypeptide I-alpha (EC 1.9.3.1) (Cytochrome A3 subunit 1-alpha).
 GN Name=ctaD; Synonyms=coi;
 OS Paracoccus denitrificans.


```

RA Capela D., Barloy-Hubler F., Couzy J., Bothe G., Ampe F., Batur J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Goffie I., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetle D., Puhler A., Purnelle B., Rampsperger U.,
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.,
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
DR EMBL; AL591785; CAC45477.1; -.
DR HSP; P33517; IM56.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0005746; C: mitochondrial electron transport chain; IEA.
DR GO; GO:0004129; F: cytochrome-c oxidase activity; IEA.
DR GO; GO:0016491; F: oxidoreductase activity; IEA.
DR GO; GO:0006118; P: electron transport; IEA.
DR GO; GO:0006810; P: transport; IEA.
DR InterPro; IPR000883; COX1.
DR Pfam; PF00115; COX1; 1.
DR PRINTS; PS01165; CYCOXIDASE1.
DR PROSITE; PS00077; COX1; 1.
KW Complete proteome; Electron transport; Heme; Oxidoreductase;
KW Respiratory chain; Transmembrane; Transport.
SQ SEQUENCE 562 AA; 62129 MW; 394144C173583612 CRC64;

Query Match
Best Local Similarity 63.3%; Score 1898.5; DB 2; Length 562;
Matches 348; Conservative 77; Mismatches 117; Indels 9; Gaps 7;

QY 3 DAAIGH-DH-HEKGGFTRFMSTNHKDIGLLYLVAAGVGFISVLFTVMRLMDPG 60
DB 13 DSHDGHADHAKHPTFFRFLSTNHKDIGLLYLVAAGVGFISVLFTVMRLMDPG 72
QY 61 VQYMCLEG-ARLIADASQTCANG--HLWNVMVTYHGLMFFVGPALFQGFNGYMLPL 117
DB 73 IQ--IFHGLAQVYGFEGDAIDGGKHMFVFTTAHALIMFFVMYPALIGFANWVPI 130
QY 118 QIGADMAFPRNNLSFWLFIAGTAGVASLAPGCGDQLGSGVGWVLYPLSTR-EGY 176
DB 131 MIGADMAFPRNNLSFWLFIAGTAGVASLAPGCGDQLGSGVGWVLYPLSTR-EGY 190
QY 177 SMDLAFVHLSGASSIMGAINMFTFLNRPAGMTLHKVPLFSWIFITAWLLILALPV 236
DB 191 AMDLAILGLHAGASSILGAINMFTFLNRPAGMTLHKVPLFSWIFITAWLLILALPV 250
QY 237 LAGAITMLTDNRNFTFFNPPAGGDPILYQHILMFFGHPEVYIILPGFIISHVSTF 296
DB 251 LAGAITMLTDNRNFTFFNPPAGGDPILYQHILMFFGHPEVYIILPGFIISHVSTF 310
QY 297 SKKPYGVLPMYVAIVGVLGVVVAHMYTVGMSLTQGSYFVLATWIAVPTGIKIFS 356
DB 311 SKRPILFGYLGAAYVAIVGVLGVVVAHMYTVGMSLTQGSYFVLATWIAVPTGIKIFS 370
QY 357 WIATWGGSVBPKSMLWAFGMFLFTYGGVTGIVLAQGLDRAYHDTYVVAHFHYMS 416
DB 371 WIATWGGSVBPKSMLWAFGMFLFTYGGVTGIVLAQGLDRAYHDTYVVAHFHYMS 430
QY 417 LGAIFAIPAGIYFMPKSGRAPPAWAKLHFWTFICANVTFFPQHFLGQGPRIID 476
DB 431 LGAIFAIPAGIYFMPKSGRAPPAWAKLHFWTFICANVTFFPQHFLGQGPRIID 490
QY 477 YPEAFALNKVSSYGAFLAFASFLFFIVFYVTLVAGRRTRPNPWGEFADTLEWTPSP 536
DB 491 YPEAFALNKVSSYGAFLAFASFLFFIVFYVTLVAGRRTRPNPWGEFADTLEWTPSP 549
QY 537 PPAHTFETLPK 547
DB 550 PPFHOWEQLEPR 560

RESULT 5
Q8UHB4 PRELIMINARY; PRT; 572 AA.
ID Q8UHB4

```

```

AC Q8UHB4;
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Cytochrome-c oxidase chain I.
GN NamesCoxA; OrderedLocusNames=Atu0768;
OC Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Dupont;
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McCelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RL Science 294:2317-2323(2001).
CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
DR EMBL; AE009044; AAL41784.1; -.
DR PIR; AB2671; AB2671.
DR HSP; P98002; IARI.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0005746; C: mitochondrial electron transport chain; IEA.
DR GO; GO:0004129; F: cytochrome-c oxidase activity; IEA.
DR GO; GO:0016491; F: oxidoreductase activity; IEA.
DR GO; GO:0006118; P: electron transport; IEA.
DR GO; GO:0006810; P: transport; IEA.
DR InterPro; IPR000883; COX1.
DR Pfam; PF00115; COX1; 1.
DR PRINTS; PS01165; CYCOXIDASE1.
DR PROSITE; PS00077; COX1; 1.
KW Complete proteome; Electron transport; Heme; Oxidoreductase;
KW Respiratory chain; Transmembrane; Transport.
SQ SEQUENCE 572 AA; 63022 MW; 1634BCF6E6005A4F CRC64;

Query Match
Best Local Similarity 63.3%; Score 1897.5; DB 2; Length 572;
Matches 350; Conservative 73; Mismatches 112; Indels 13; Gaps 6;

QY 7 HGHDEHKQGFTRFMSTNHKDIGLLYLVAAGVGFISVLFTVMRLMDPGVYMCL 66
DB 29 HSHDGHKPGFFARWFLSTNHKDIGLLYLVAAGVGFISVLFTVMRLMDPGVYMCL 86
QY 67 EG-ARLIADASQTCANG--HLWNVMVTYHGLMFFVGPALFQGFNGYMLQIGAPD 123
DB 87 HGLASVMYGFEGDAIDGGKHMFVFTTAHALIMFFVMYPALIGFANWIPIMIGAPD 146
QY 124 MAPFRNNLSFWLFIAGTAGVASLAPGCGDQLGSGVGWVLYPLSTR-EGYSVDLAI 182
DB 147 MAPFRNNLSFWLFIAGTAGVASLAPGCGDQLGSGVGWVLYPLSTR-EGYSVDLAI 206
QY 183 FAVHLSGASSIMGAINMFTFLNRPAGMTLHKVPLFSWIFITAWLLILALFVLGAGIT 242
DB 207 FSLHVAGASSILGAINMFTFLNRPAGMTLHKVPLFSWIFITAWLLILALFVLGAGIT 266
QY 243 MLLTDRNFTFFNPPAGGDPILYQHILMFFGHPEVYIILPGFIISHVSTFSSKPKVF 302
DB 267 MLLTDRNFTFFNPPAGGDPILYQHILMFFGHPEVYIILPGFIISHVSTFSSKPKVF 326
QY 303 GYLPMYVAIVGVLGVVVAHMYTVGMSLTQGSYFVLATWIAVPTGIKIFSIAWTMW 362
DB 327 GYLPMYVAIVGVLGVVVAHMYTVGMSLTQGSYFVLATWIAVPTGIKIFSIAWTMW 386

```

QY 363 GGSVEFKSPMLWAFGPMFLFTVGGTIGIVLAQAGLDRAHYDHYVVAHPHYVMSLGAIFA 422
 DB 387 GGSLSLTESTPMVAIGFIPLFTVGGTIGVQLANAGLDRLSDHYVVAHPHYVLSLGAIFA 446
 QY 423 IFAGIYFYPKFSGRAPFPAWAKLHFWTFIFIGANVTFFQHFGLGQGMRRRIIDYPEAFA 482
 DB 447 IFAGIYFYPKFSGRAPFPAWAKLHFWTFIFIGANVTFFQHFGLGQGMRRRIIDYPEAFA 506
 QY 483 LMKVSSYGAFL--AFASGLFFIVFVTVLVAQRETRPNPFGSPADTLEWTLSPPPA 539
 DB 507 GNNVSSYSYISAVAVGIFLP----GVNEAFKRIAGNNPAGEGATTLEWTLSPPPA 562
 QY 540 HTPETLPK 547
 DB 563 HOWEQLPR 570
 RESULT 6
 Q8G263 PRELIMINARY; PRT; 552 AA.
 AC Q8G263;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DE 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Cytochrome c oxidase, subunit I (EC 1.9.3.1).
 GN Names:coxI; OrderedLocusNames=BR0468;
 OS Brucella suis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 NCBI_TaxID=29461;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1330 / Biovar 1;
 RX MEDLINE=2224741; PubMed=12271122; DOI=10.1073/pnas.192339099;
 RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
 RA Read T.D., Dodson R.J., Umayam L.A., Brinkac L.M., Bearan M.J.,
 RA Daugherty S.C., DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
 RA Nelson W.C., Avodeji B., Kraul M., Shetty J., Malek J.A.,
 RA Van Aken S.E., Riedmuller S., Tettelin H., Gill S.R., White O.,
 RA Salzberg S.B., Hoover D.L., Linder L.H., Hailing S.M., Boyle S.M.,
 RA Fraser C.M.;
 RT "The Brucella suis genome reveals fundamental similarities between
 RT animal and plant pathogens and symbionts."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
 CC -1- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
 DR EMBL; AE014356; AN29411.1; -;
 DR HSSP; P33517; IM56.
 DR TIGR; BR0468; -;
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000883; COXI.
 DR Pfam; PF00115; COXI; 1.
 DR PRINTS; PR01165; CYCOXIDASEI.
 DR PROSITE; PS00077; COXI; 1.
 DR Complete proteome; Electron transport; Heme; Oxidoreductase;
 KW Respiratory chain; Transmembrane; Transport
 SQ SEQUENCE 552 AA; 60723 MW; 072B846C47D628C RC64;
 Query Match 63.0%; Score 1891.5; DB 2; Length 552;
 Best Local Similarity 63.0%; Pred. No. 2.9e-114;
 Matches 351; Conservative 76; Mismatches 113; Indels 17; Gaps 8;
 QY 1 MADAATH--GPHDHEKQGFFFWFSTNKHDTGLLYVAAGVGFISLVFVYMLELM 57
 DB 1 MAGTAAHEGAHDHDKPHG--VWRVWYSTNKHDTGLLYLIPAIAGLIGALSINRAELQ 59
 QY 58 DPGVQVYKLEG-ARLI--ADASQCTCTANGHLMNVMYVTHGILMFFVGIPALFGGFGNLY 114
 DB 60 EPGIQ--IFHGLAQMVGVGSDAALDAGKMFNVFTSAHALVMIFVMVFPALIGGFANWM 117

QY 115 MPLQIGAPDMAPFRMNNLSFWLFIAGTAMGVSALFAPGGDQGLSGVGVLYPLPLSTR-E 173
 DB 118 VPLMIGAPDMAPFRMNNLSFWLFIAGTAMGVSALFAPGGDQGLSGVGVLYPLPLSTR-E 177
 QY 174 AGYSMDLAI FAVHLSGASSIMGAINMITTFLNRAQGMTHKVLPSWSIFITAMILLIA 233
 DB 178 PGPAVDFAILAIHISGASSILGAINFITITLNNRAPGMTHKMLPFAVAVLGFATALLIS 237
 QY 234 LPVLAGATMTLLDRNFGTTFNPNAGGGDPIYLQHLILWFFGHPEVYIIILPGFGIISHVV 293
 DB 238 LPVLAGATMTLLDRNFGTTFNPNAGGGDPIYLQHLILWFFGHPEVYIIILPGFGIISHVV 297
 QY 294 STFSKPEVFGYLPVYANVAIGVLGVVVAHMYTVGMSLTQOSYFMLATMVIJAVTGK 353
 DB 298 STFSKPEVFGYLPVYANVAIGVLGVVVAHMYTVGMSLTQOSYFMLATMVIJAVTGK 357
 QY 354 IFSWIATMWSGVSEFKSPMLWAFGPMFLFTVGGTIGIVLAQAGLDRAHYDHYVVAHPHY 413
 DB 358 IFSWIATMWSGVSEFKSPMLWAFGPMFLFTVGGTIGIVLAQAGLDRAHYDHYVVAHPHY 417
 QY 414 VMSLGAIFAIFAGIYFYPKFSGRAPFPAWAKLHFWTFIFIGANVTFFQHFGLGQGMRR 473
 DB 418 VLSLGAIFAIFAGIYFYPKFSGRAPFPAWAKLHFWTFIFIGANVTFFQHFGLGQGMRR 477
 QY 474 YIDYPEAFALWNVSSYGAFLFAFASGLFFIVFVTVL--VAGRRETRPNPFGSPADTLE 530
 DB 478 YIDYPEAFALWNVSSYGAFLFAFASGLFFIVFVTVL--VAGRRETRPNPFGSPADTLE 533
 QY 531 WTLSPPPAHTFETLPK 547
 DB 534 WQLSSPPPPHQBQLPR 550
 RESULT 7
 Q8YFQ5 PRELIMINARY; PRT; 552 AA.
 AC Q8YFQ5;
 DT 01-MAR-2002 (Tremblrel. 20, Created)
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
 DE 01-MAR-2004 (Tremblrel. 25, Last annotation update)
 DE CYTOCHROME C OXIDASE POLYPEPTIDE I (EC 1.9.3.1).
 GN OrderedLocusNames=BME11465;
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 NCBI_TaxID=29459;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=16M / ATCC 23456 / Biotype 1;
 RX MEDLINE=20020109; PubMed=11756688; DOI=10.1073/pnas.221575398;
 RA DelVecchio V.G., Kapral V., Redkar R.J., Patra G., Mijer C., Los T.,
 RA Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,
 RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Gotsman E.,
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
 RA Haselkorn R., Kyrides N.C., Overbeek R.;
 RT "The genome sequence of the facultative intracellular pathogen
 RT Brucella melitensis."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 CC -1- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
 DR EMBL; AE009593; AL52646.1; -;
 DR FIR; AC3435; AC3435.
 DR HSSP; P33517; IM56.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000883; COXI.
 DR Pfam; PF00115; COXI; 1.
 DR PRINTS; PR01165; CYCOXIDASEI.
 DR PROSITE; PS00077; COXI; 1.
 DR Complete proteome; Electron transport; Heme; Oxidoreductase;
 KW Respiratory chain; Transmembrane; Transport
 SQ SEQUENCE 552 AA; 60723 MW; 072B846C47D628C RC64;
 Query Match 63.0%; Score 1891.5; DB 2; Length 552;
 Best Local Similarity 63.0%; Pred. No. 2.9e-114;
 Matches 351; Conservative 76; Mismatches 113; Indels 17; Gaps 8;
 QY 1 MADAATH--GPHDHEKQGFFFWFSTNKHDTGLLYVAAGVGFISLVFVYMLELM 57
 DB 1 MAGTAAHEGAHDHDKPHG--VWRVWYSTNKHDTGLLYLIPAIAGLIGALSINRAELQ 59
 QY 58 DPGVQVYKLEG-ARLI--ADASQCTCTANGHLMNVMYVTHGILMFFVGIPALFGGFGNLY 114
 DB 60 EPGIQ--IFHGLAQMVGVGSDAALDAGKMFNVFTSAHALVMIFVMVFPALIGGFANWM 117

KW Complete proteome; Electron transport; Heme; Oxidoreductase;
 KW Respiratory chain; Transmembrane; Transport.
 SQ SEQUENCE 552 AA; 60768 MW; 022505D26753E769 CRC64;

Query Match 63.0%; Score 1890.5; DB 2; Length 552;
 Best Local Similarity 62.8%; Pred. No. 3.4e-114;
 Matches 350; Conservative 77; Mismatches 113; Indels 17; Gaps 8;

Qy 1 MADAAIH---CHDHEKQGFTRFMSTNKHDKIGLLVLAAGVGVFTSVLFTVYMRLELM 57
 Db 1 MAGTAAHEHGADHDKPHG-WVRVWYTNKHDKIGTLTILFALITAGTGGALSTAMRAELQ 59

Qy 58 DPGVQYMCLEG-ARLI--ADASQCTCTANGHLNWNVYTHGILMFFVIGIPALFGGFGNYL 114
 Db 60 EPGIQ--IFHGLAQVYGVGEADALDACKMFNVFTSAHALVFMFFVMVMPALLIGFANWM 117

Qy 115 MPLOIGAPDMAPPRNNLSFWLFTAGTAMGVASLFAFGGQDLGSGVGWVLYPLSTR-E 173
 Db 118 VPLMIGAPDMAPPRMNNISFWLLPPALTLISLNFPGPAGGPGGGGWTLYPPFSTSGQ 177

Qy 174 AGYSNDLAIFAVHLSGASSIMGAINMITFLNMRAPGNTLHKVPLFSWIFITAWLILLA 233
 Db 178 PGPAVDFAILAIHISGASSILGAINFTILNMRAPGNTLHKMPLFAWAVLITAFILLLS 237

Qy 234 LPVLAGATMLLTDNRNFTFFNPAGGDPILYQHILWFFGHPDEVYIIILPGFGIISHV 293
 Db 238 LPVLAGGITMLLTDNRNFTFFNPAGGDPILYQHLEWFFGHPDEVYIIILPGFGIISHV 297

Qy 294 STFSKPKVFGYLPVMAVVAITGLGVVVAHMTVGMSLTQCSYFVMTATWIAVPTGIK 353
 Db 298 STFSKRPFGYLGWAVVAITGLGVVVAHMTVGLSTQRYFVPTATWIAVPTGIK 357

Qy 354 IFSWIATWGGSVFEPKSPMLWAFGFMFLFTVGGVGTIVLAQAGLDRAVHDYVVAHFY 413
 Db 358 IFSWIATWGGSVISPRAPMLWAGFIPLFTTGGVGTIVLAQAGLDRAVHDYVVAHFY 417

Qy 414 VMSLGAIFAFAIGIYFVMPKSGRAFPWAAKLHPWTFPIGANTFFPOHFLGQGMPPR 473
 Db 418 VLSLGAIFAFAIGIYFVMPKSGRYNNEFTAKLHPWTFVGNLVNFFPOHFLGQGMPPR 477

Qy 474 YIDYEAALNKKVSSYGAFLAFASFLFFVIFVYTL---VAGRETRPNPWFGEADTLE 530
 Db 478 YIDYDPAVAGNWNISYSGYSISG---FAVLIFLYNVEFAAKKEAGANPGDGATTLE 533

Qy 531 WTLPSPPPAHTPETLPK 547
 Db 534 WOLSSPPPPHQQEQLPR 550

RESULT 8
 COX1 RHILE STANDARD; PRT; 538 AA.
 AC Q08855;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Cytochrome c oxidase polypeptide I (EC 1.9.3.1) (Cytochrome AA3 subunit 1).
 GN Name=ctaD; Synonyms=coxA;
 OS Rhizobium leguminosarum.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
 OX NCBI_TaxID=384;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94161496; PubMed=8117073;
 RA Gabel C., Bittinger M.A., Maier R.J.;
 RT "Cytochrome aa3 gene regulation in members of the family Rhizobiaceae:
 RT comparison of copper and oxygen effects in Bradyrhizobium japonicum
 RT and Rhizobium tropici";
 RL Appl. Environ. Microbiol. 60:141-148 (1994).
 CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
 chain that catalyzes the reduction of oxygen to water. Subunits 1-

3 form the functional core of the enzyme complex. CO I is the catalytic subunit of the enzyme. Electrons originating in cytochrome c are transferred via the copper A center of subunit 2 and heme A of subunit 1 to the bimetallic center formed by heme A3 and copper B.
 -!- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome c + 2 H(2)O.
 -!- PATHWAY: Respiratory chain; terminal step.
 -!- SUBCELLULAR LOCATION: Integral membrane protein.
 -!- DEVELOPMENTAL STAGE: Free in soil (not as bacteroid).
 -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

EMBL; X74341; CAAS2388.1; -.
 PIR; S36424; S36424.
 HSP; P98002; IQLE.
 Inter-Pro; IPR000883; COX1.
 Pfam; PF00115; COX1; 1.
 PRINTS; PR01165; CYCOXIDASE1.
 PROSITE; PS00855; COX1; 1.
 PROSITE; PS00077; COX1_CUB; 1.
 DR Copper; Heme; Oxidoreductase; Respiratory chain; Transmembrane.
 FT TRANSMEM 37 57 Potential.
 FT TRANSMEM 84 104 Potential.
 FT TRANSMEM 121 141 Potential.
 FT TRANSMEM 170 190 Potential.
 FT TRANSMEM 208 228 Potential.
 FT TRANSMEM 259 279 Potential.
 FT TRANSMEM 291 311 Potential.
 FT TRANSMEM 329 349 Potential.
 FT TRANSMEM 362 382 Potential.
 FT TRANSMEM 401 421 Potential.
 FT TRANSMEM 438 458 Potential.
 FT TRANSMEM 476 496 Potential.
 FT METAL 82 82 Iron (heme A axial ligand) (Probable).
 FT METAL 265 265 Copper B (Probable).
 FT METAL 269 269 Copper B (Probable).
 FT METAL 314 314 Copper B (Probable).
 FT METAL 315 315 Copper B (Probable).
 FT METAL 400 400 Iron (heme A3 axial ligand) (Probable).
 FT METAL 402 402 Iron (heme A axial ligand) (Probable).
 FT METAL 402 402 1'-histidyl-3'-tyrosine (His-Tyr) (By similarity).
 FT CROSSLNK 265 269
 SQ SEQUENCE 538 AA; 59133 MW; 740CBFE6C902D9BF CRC64;

Query Match 62.5%; Score 1873.5; DB 1; Length 538;
 Best Local Similarity 64.2%; Pred. No. 4.1e-113;
 Matches 348; Conservative 63; Mismatches 110; Indels 21; Gaps 4;

Qy 7 HGHDEHEKQGFTRFMSTNKHDKIGLLVLAAGVGVFTSVLFTVYMRLEMDGQVQMCL 66
 Db 15 HGHDEHAHP---TCWRRTNKHDKIGTLTILFALITAGVIGAMSLAIRAELMYPGVEYF-- 69

Qy 67 EGARLIADASQCTCTANGHLNWNVYTHGILMFFVIGIPALFGGFGNYLMPLOIGAPDMAP 126
 Db 70 -----HNTHLYNVFTSHGVIMIFMVPANMIGGFGNFWFLMTGAPDMAP 115

Qy 127 PRMNLISFWLFIAGTAMGVASLFAFGGQDLGSGVGWVLYPLSTR-EAGYSMDLAIFAV 185
 Db 116 PRMNLISFWLIPASFGILLMSTFVEGPGANGAGAGAGWMTVPLSSSGHPGPAVDLAIFSL 175

Qy 186 HLCGASSIMGAINMITFLNMRAPGNTLHKVPLFSWIFITAWLILALPVLGATIMLL 245
 Db 176 HIACASSILGAINFITILNMRAPGNTLHKMPLFANWSVLITAFILLLSLPLVLAGATIMLL 235

Qy 246 TDRNFGITTFNPAGGDPILYQHILWFFGHPDEVYIIILPGFGIISHVSVTSKPKVFGYL 305

Db 236 TDRNFGTTFAPSGGGDPLLYQHLPFFGPHGVVILILPGFGMISHIISTFSRKPVEGYI 295
 QY 306 PMYAMVAIGVLGVVWVAHMYTVGSLTQOSYFMLATVIAVPTGIGKIPSMIATWGGG 365
 Db 296 GMVYAAAIIGLGVVWVAHMYIVGMDLDEATFVSATMIIAVPTGIGKIPSMIATWGGG 355
 QY 366 VFKSPMLAFGFMELFTVGGVGTGIVLAQGLDRAYHDYVYVVAHPHYVMSLGAIFAIFA 425
 Db 356 IEPATPMLALATIFELFTVGGVGTGIVLAQGLDRAYHDYVYVVAHPHYVMSLGAIFAIFA 415
 QY 426 GIYFVMPKSGRAPPWAALKHWTFFIGANVTFFPHQHLGRCQMPRRVIDIDPEAPALWN 485
 Db 416 GWIYFPMXGMYNETLAEHFWLIFIGVNLFFPEHFLGSGMPRRVIDIDPEAFAGWN 475
 QY 486 KVSYSYGAFLAFASFLFFIVFVTVLAGRRTRPNPWSFADTLEWTLSPSPPAITFETL 545
 Db 476 LVSSIGSYISGFSVLLFIYCVYDAFAKNVPVGDNPWAGATLLEWTLSPSPVHEFVYL 534
 QY 546 PK 547
 Db 535 PR 536

RESULT 9

Q9A300 PRELIMINARY; PRT; 552 AA.
 AC Q9A300; 2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Cytochrome c oxidase, subunit I.
 GN OrderedLocusNames=CC3406;
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
 OC Caulobacteraceae; Caulobacter.
 OX NCBI_TaxID=155892;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 19089 / CB15;
 RX MEDLINE=21173698; PubMed=11259647; DOI=10.1073/pnas.06109298;
 RA Niernan W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H.M., Shetty J.,
 RA Berry K.J., Utterback T.R., Tran K., Wolf A.M., Vamathevan J.J.,
 RA Emolaeva M.D., White O., Salzberg S.L., Venter J.C., Shapiro L.,
 RA Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 CC -/- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
 DR EMBL: AB006001; AAK25368.1; -;
 DR FTR: D87671; D87671.
 DR HSP: P00396; 20CC.
 DR TIGR: CC3406; -;
 DR GO: GO:0016021; C:integral to membrane; IEA.
 DR GO: GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO: GO:0004129; F:cytochrome-c oxidase activity; IEA.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0006118; P:electron transport; IEA.
 DR GO: GO:0006810; P:transpor; IEA.
 DR InterPro: IPR000883; COX1.
 DR Pfam: PF00115; COX1; 1.
 DR PRINTS: PR01165; CYCOXIDASE1.
 DR PROSITE: PS00077; COX1; 1.
 DR Complete proteome; Electron transport; Heme; Oxidoreductase;
 KW Respiratory chain; Transmembrane; Transport.
 SQ SEQUENCE 552 AA; 6101 MW; E82191E54206AACS CRC64;

Query Match 61.6%; Score 1847.5; DB 2; Length 552;
 Best Local Similarity 61.6%; Pred. No. 26-111;
 Matches 341; Conservative 70; Mismatches 136; Indels 7; Gaps 4;

QY 2 ADAATHGHDEHKQGFTRFMFSTNHDKDIGLLVLAAGVVGFTSVLFTVVMRLMLDPGV 61
 Db 5 ADTHDAHDDDEKPPFARFWFFSNNHDKDIGLLYILPALMAGIVGGALSGLIRNELMXPAGI 64
 QY 62 QYMCLEGARLIADASQCTTANGHLWNNVYVHGLMFFVGIIPALFGGFGNYLMLPQIGA 121
 Db 65 QVPSDTGML-----AQMGIFKKGHCYNAAVTAHALIMIFFMVMIDAMIGGFGNFWFPIMIGA 120
 QY 122 PDMAFPPRNLSFWLFIAGTAMGVASLFAFGDQGLGSGVGVLYPPLST-REAGYSMDL 180
 Db 121 PDMAFPPRNLSFWLFIAGTAMGVASLFAFGDQGLGSGVGVLYPPLST-REAGYSMDL 179
 QY 181 AIFAVHLSGASSIMGAINMITTFNNRPAQVTLHKVPLFSWISFITAFLILLALPVLAGA 240
 Db 180 AILSLHLAGASSILGAINFITTFNMEAPGTLHRMPLFVWSVLITAFLLLSLPVLAGA 239
 QY 241 ITMLLTDRNFGTTFENPAGGDDPILYOHILWFRGHPEVYIILPGGCIISHVYSTSKKP 300
 Db 240 ITMLLTDRNFGTTFENPAGGDDPILYOHILWFRGHPEVYIILPGGCIISHVYSTSKKP 299
 QY 301 VFGYLPVYAMVAIGVLGVVWVAHMYTVGSLTQOSYFMLATVIAVPTGIGKIPSMIAT 360
 Db 300 VFGYLPVYAMVAIGVLGVVWVAHMYTVGSLTQOSYFMLATVIAVPTGIGKIPSMIAT 359
 QY 361 MWGGSVEFKSPMLWAFQFMFLFTVGGVGTGIVLAQGLDRAYHDYVYVVAHPHYVMSLGA 420
 Db 360 MWGGSVEFKSPMLWAFQFMFLFTVGGVGTGIVLAQGLDRAYHDYVYVVAHPHYVMSLGA 419
 QY 421 FAIFAGIYFVMPKSGRAPPWAALKHWTFFIGANVTFFPHQHLGRCQMPRRVIDIDPEA 480
 Db 420 FAIFAGIYFVMPKSGRAPPWAALKHWTFFIGANVTFFPHQHLGRCQMPRRVIDIDPEA 479
 QY 481 PALMKNVSSYGAFLAFASFLFFIVFVTVLAGRRTRPNPWSFADTLEWTLSPSPPAH 540
 Db 480 FTLWNVYSTVGYMTIVVGVGVFVMLVIEAIR-RRKAEANPWEAGATLLEWTLSPSPPH 538
 QY 541 TFEETLPKESDWDKH 554
 Db 539 QFSEFPVKGDDHH 552

RESULT 10

Q985X1 PRELIMINARY; PRT; 550 AA.
 AC Q985X1; 2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Cytochrome c oxidase subunit I.
 GN OrderedLocusNames=mlr7491;
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 Mesorhizobium loti.";
 RL DNA Res. 7:331-338(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF303099;
 RX MEDLINE=21082936; PubMed=11214974;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,

RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti (supplement).";
RL DNA Res. 7:381-406(2000).
CC -1- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
DR EMBL; AP003011; BAB53582.1; -;
DR HSSP; P33317; 1M56;
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0005746; C: mitochondrial electron transport chain; IEA.
DR GO; GO:0004129; F: cytochrome-c oxidase activity; IEA.
DR GO; GO:0016491; F: oxidoreductase activity; IEA.
DR GO; GO:0006118; P: electron transport; IEA.
DR GO; GO:0006810; P: transport; IEA.
DR InterPro; IPR000883; COX1.
DR Pfam; PF00115; COX1; 1.
DR PRINTS; PR01165; CYCOXIDASE1.
DR PROSITE; PS00077; COX1; 1.
DR KW Complete proteome; Electron transport; Heme; Oxidoreductase;
KW Respiratory chain; Transmembrane; Transport.
SQ SEQUENCE 550 AA; 60241 MW; B3477BFA1B1AE67D CRC64;

Query Match 61.6%; Score 1846.5; DB 2; Length 550;
Best Local Similarity 61.4%; Pred. No. 2.3e-111;
Matches 341; Conservative 79; Mismatches 120; Indels 15; Gaps 8;

QY 1 MADAAIH-GDHHEKQGFTRWFSTNKHDTGLLYVAAGVGVGISVLFTVYMRLELMDP 59
Db 1 MADAAAHGHCHKPHVG-WVRVYSTNKHDTGLYLFAIMAGILGALSVAIRVELOEP 59
QY 60 GVQVMCLSEG-ARLIADASQCTTANG--HLNVMVYTHGILMMFPVGPALFGGNGVIMP 116
Db 60 GIQ--IFSGLAQVMYGMNGDAIDGKSMYNAFAHAHALIMFFVMFALIGGFANMVP 117
QY 117 LOIGAPDMAPRMNNLSFWLFIAGTAMGVASLFAFGGQGLGSGVGWLYVPLSTR-EAG 175
Db 118 IMIGAPDMAPRMNNISFWLLPPAPILLLTSNFPVPSAPAGVGCGWTLVPLSTSGOP 177
QY 176 YSMDLAI FAVHLSGASSINGAINMITTFNMRAPGMLTKVPLFSWSIFITAWLILALP 235
Db 178 PAMDALILSIHAGASSILGAINFITTFNMRAPGMLTKVPLFSWSVLTAFLLLSLP 237
QY 236 VLAGAITMLLTDNRNFTTFPAGGDPILYQHILWFFGHGHEVYIILPGGIIISHVYST 295
Db 238 VLAGGITMLLTDNRNFTTFPAGGDPILYQHILWFFGHGHEVYIILPGGIIISHVYST 297
QY 296 FSKKPVFGLPMYAWAIVGLGVVVAHMYTVGMSLTQOSYFMLATMWIAVPTGIKIF 355
Db 298 FSKKPVFGLMAYANVAIVAGFVVAHMYTVGMSLTQGLSDTQRYFVFAIMVIAVPTGVKIF 357
QY 356 SWIATMWGSEVFKSPMLWAFGMFLFTVGGVTGIVLAQAGLDRAHYHDTYVVAHFHYVM 415
Db 358 SWIATMWGSGISFKTFLMALGFILFTIGTGGTVGVLNAGLDRLSHDITYVIAHFHYVL 417
QY 416 SIQALFAIPAGYFYFMPKESGRAPFEWAKLHFMTFFRIGANVTFPPQHLGQGMPERYI 475
Db 418 SIQAVFAIPAGYFYFMPKMGYINSEVANTHFVTFVGVNLIFFPQHLGAGMPRTI 477
QY 476 DYPEAFALNKNVSSYCAFLAFASFLFFIVIFYTLVAGRRRETR---PNPWGEFADTLWE 532
Db 478 DYPEAFAGNYSYSGSYISAVG---VAIFLYGVFEAFQKKRIAGANPWGEATLLEWQ 533
QY 533 LPSPPPAHTFTLPK 547
Db 534 LPSPPPFHOWELPK 548

RESULT 11
Q6NBJ6 PRELIMINARY; PRT; 541 AA.
AC Q6NBJ6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytochrome c oxidase subunit I
DE Name=COXA; OrderedAccession=RPA0832;
OS Rhodospseudomonas palustris.
CC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
CC Bradyrhizobiaceae; Rhodospseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=GA009 / ATCC BAA-98;
RX PubMed=14704707; DOI=10.1038/nbt933;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA Harrison F.H., Gibson J., Harwood C.S.;
RA "Complete genome sequence of the metabolically versatile
RT photosynthetic bacterium Rhodospseudomonas palustris.";
RL Nat. Biotechnol. 22:55-61(2004).
CC -1- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
DR EMBL; BX572595; CAE26276.1; -;
DR InterPro; IPR000883; COX1.
DR Pfam; PF00115; COX1; 1.
DR PRINTS; PR01165; CYCOXIDASE1.
DR PROSITE; PS00077; COX1; 1.
DR KW Complete proteome; Electron transport; Heme; Oxidoreductase;
KW Respiratory chain; Transmembrane; Transport.
SQ SEQUENCE 541 AA; 59662 MW; 4C827D1C780FC41B CRC64;

Query Match 61.4%; Score 1842; DB 2; Length 541;
Best Local Similarity 63.1%; Pred. No. 4.5e-111;
Matches 346; Conservative 68; Mismatches 106; Indels 28; Gaps 7;

QY 7 HGHDDH---EKOGFFTRWFMSNHNKDIGLLYLVAAGVGVISVLFTVYMRLELMDPGVQY 63
Db 13 HAHDHDAHAPTC-WRRYLYSTNHNKDIGTMYLVFAVMAGVIGASIAIRVELMFFGIQI 71
QY 64 MCEGARLIADASQCTTANGHLNVMVYTHGILMMFFVGPALFGGNGVLMFLQIGAPD 123
Db 72 F-----HNTHYVNFVSHGLNIFVWNPAMVIGFNGVFWFLMIGAPD 115
QY 124 MAPPRMNNISFWLFIAGTAMGVASLFAFGGQGLGSGVGWLYVPLST-REAGYSMDLAI 182
Db 116 MAPPRMNNISFWLLPAAFALLIISTFVEGEPSSNGVGAGTWYAPLSTIGHPGPAVDFAI 175
QY 183 FAVHLSGASSINGAINMITTFNMRAPGMLTKVPLFSWSIFITAWLILALPVLGAI 242
Db 176 LALHLGASSILGAINFITTFNMRAPGMLTKVPLFSWSLITAFLLLSLVLGAI 235
QY 243 MLLTDNRNFTTFPAGGDPILYQHILWFFGHGHEVYIILPGGIIISHVYSTFSKKPV 302
Db 236 MLLTDNRNFTTFPAGGDPILYQHILWFFGHGHEVYIILPGGIIISHVYSTFSKKPV 295
QY 303 GYLPWYAMVAIVGLGVVVAHMYTVGMSLTQOSYFMLATMWIAVPTGIKIFSWIATMW 362
Db 296 GYLGMAYAMVAIVGIVGVVVAHMYTVGMSLTQAYFVAATMWIAVPTGVKVFESWIATMW 355
QY 363 GGSVEFKSPMLWAFGMFLFTVGGVTGIVLAQAGLDRAHYHDTYVVAHFHYVMVSLGAIFA 422
Db 356 GGSIEFPKTPMLWAVGFIPLFTVGGVTGIVLANAGVRLVQDTYVVAHFHYVLSLGAIFA 415
QY 423 IFAGIYFMPKESGRAPFEWAKLHFMTFFRIGANVTFPPQHLGQGMPERYI DYPEAFA 482
Db 416 IFAGIYFMPKMGYINSEVANTHFVTFVGVNLIFFPQHLGAGMPRYDYDPDAPA 475
QY 483 LWNKVSSYGAFLAPASFLFFIVIFY--TLVAGRRRE--TRPNWGEFADTLLEWLPSPPPA 539
Db 476 GWNLVSSIGSYISG-----FAVLVFLVGMTLAFIRKERAADNPWGAGATTLLEWLSGPPPF 531
QY 540 HTPETLPK 547
Db 532 HQFEVLQ 539

RESULT 12
CAE26276 PRELIMINARY; PRT; 541 AA.
ID Q59631 AC Q59631; PRT; 539 AA.
CAE26276; PRT; 541 AA.
DT 02-MAR-2004 (TREMBlrel. 27, Created)
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
DE Cytochrome c oxidase subunit 1.
DE Cytochrome c oxidase subunit 1 (EC 1.9.3.1).
GN COXA OR RPA0832.
OS Rhodospseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodospseudomonas.
OX NCBI_TaxID=1076;
RN [1]_TaxID=1076;
SEQUENCE FROM N.A.
RP STRAIN=CGA009 / ATCC BAA-98;
RC PubMed=14704707;
RX Larimer F.W., Chain P., Hauser L., Lamerdin J., Malfatti S., Do L., Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R., Gibson J.L., Hanson T.E., Bobst C., Torres J.L., Peres C., Harrison F.H., Gibson J., Harwood C.S.;
RA "Complete genome sequence of the metabolically versatile phototrophic bacterium Rhodospseudomonas palustris."
RT Nat. Biotechnol. 22:55-61(2004).
DR EMBL; BX572595; CAE26276.1; -.
SQ SEQUENCE 541 AA; 59662 MW; 4C927D1C780FC41B CRC64;
Query Match 61.4%; Score 1842; DB 2; Length 541;
Best Local Similarity 63.1%; Pred. No. 4.5e-111;
Matches 346; Conservative 68; Mismatches 106; Indels 28; Gaps 7;
QY 7 HGHDRH---EKQGFTRFMSTNKHDIIGLLYLVAAGVGVGFSVLFTVMRLMLMDPGVOY 63
DB 13 HAHDHNAHPTG-WRRYLSTNKHDIIGTMYLVAVWAGVIGAMSAIRVELVFQIQI 71
QY 64 MCEGARLIADASQTCANGHLNNVMVTHGILMMFFVGPALFGGNYMLPQIGAPD 123
DB 72 F-----HNTYVNFVTSHGLIMIFFWVMPAMITGGFNWVFLMIGAPD 115
QY 124 MAPFRNNLSFWLFTAGTAMGVASLFPAGDQGLSGGVWLVYPLST-REAGVSMDLAI 182
DB 116 MAPFRNNLSFWLPAAPALLIISTFVEGESSNGVGAGMTYAPLSTIGHPGAPVDFAI 175
QY 183 FAVHL-SGASSINGAINMITTFELNRPAGMTLHKVPLFSWGFITAWLILALFLAGAIT 242
DB 176 LALHLAGASSILGAINFTTFINRPAQMTLHKMPLFWMSVLITAPLILLSLPVLGAI 235
QY 243 MLLTDNRNFTFFNPAGGGDILQHLWFGHEDEVIIILPGGIIISHVSTFSKKPVF 302
DB 236 MLLTDNRNFTFFSAEGCGDPLFLQHLWFGHEDEVIIILPGGIIISHVSTFSKKPVF 295
QY 303 GYLPVMVAMVAIGVLGVVAAHMYTVGMSLTQOSVFWLATVIAVPTGKIFSWATMW 362
DB 296 GYLGNAYANVAIGVLGVVAAHMYTVGMSLTQAYFVAATVIAVPTGKIFSWATMW 355
QY 363 GGSVEFKSPMLWAGFMPLFTVGGVTGVLQAQGLDRAYHDTYVVAHFYVMSLGAIFA 422
DB 356 GGSIEFKTPEMLWAGVIFLFTVGGVTGVLNAGVDEVLQDTYVVAHFYVMSLGAIFA 415
QY 423 IFAGYFVWPKFSGRAPPEAAKLHFTFTFIGNVTFPPQHFGLGRCQMPRYDYPEAFA 482
DB 416 IFAGYVWFPPQNTGYNNETGKHLFWLTFILGNWVFPQHFGLSGMPRYDYDFDFA 475
QY 483 LWNKSSYCAFLAFASFLFTFIVY--TLVAGRR--TRNPWGEFADTLLEWLPSPPPA 539
DB 476 GWNLYSSIGSVISG----FAVLVFLYGLTWLAFIRKERAADNPWAGAGATTLEWLSPPPP 531
QY 540 HTEETLPK 547
DB 532 HQFVLPPQ 539

DB 344 KVFSWIATMNGSIEFKTPTMLFAIGFLFTVGGVTGVVLANAGVDRVMOETYYVIAHFH 403
 QY 413 YVMSLGAIFALFAGFYFYMRFSGRAPEMAAKLHFWTFIFIGANVTFFPQHFLGRQGMPR 472
 DB 404 YVMGIAVFAIFSGWYWFPMWSGYMNETIGKLFHFWMFIGNILFFPQHFLGLQGMWR 463
 QY 473 RYIDYPEAFALWNVSSYGAFLAFASPLFFIVFYVTLV----AGRETNPWGEFADTL 529
 DB 464 RMVDYDPAFAGWNETSSYGAFITG----FGVIIFLYGLVDFAFARKQAADNPWGEGATTL 519
 QY 530 EWTLPSPPPPAHTEFTLPK 547
 DB 520 EWTLPSPPPPFHQTLPK 537

RESULT 14
 COX1 BRAJA
 ID COX1 BRAJA STANDARD; PRT; 541 AA.
 AC P31833;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Cytochrome c oxidase polypeptide I (EC 1.9.3.1) (Cytochrome AA3 subunit 1).
 GN Name=ctaD; Synonyms=coxA; OrderedLocusNames=blr1171;
 OS Bradyrhizobium japonicum.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Bradyrhizobiaceae; Bradyrhizobium.
 OX NCBI_TaxID=375;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LO;
 RX MEDLINE=91045095; PubMed=2172930;
 RA Gabel C., Maier R.J.;
 RT "Nucleotide sequence of the coxA gene encoding subunit I of cytochrome a3 of Bradyrhizobium japonicum.";
 RL Nucleic Acids Res. 18:6143-6143(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=USDA 110spc4;
 RX MEDLINE=91211625; PubMed=1965217;
 RA Bott M., Bolliger M., Hennecke H.;
 RT "Genetic analysis of the cytochrome c-a3 branch of the Bradyrhizobium japonicum respiratory chain.";
 RL Mol. Microbiol. 4:2147-2157(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=USDA 110spc4;
 RX MEDLINE=22484998; PubMed=12597275;
 RA Rossmann R., Loferer H., Rossi P., Hennecke H.;
 RT "Factors involved in biogenesis of active cytochrome a3 encoded by the coxAEGC gene cluster from Bradyrhizobium japonicum.";
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=USDA 110spc4;
 RX Mueller P.;
 RT "Extended sequencing of a DNA fragment of B.japonicum adjacent to the cox operon.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=USDA 110;
 RX MEDLINE=22484998; PubMed=12597275;
 RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
 RA Sasanoto S., Watanabe A., Idesawa K., Iriquchi M., Kawashima K.,
 RA Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
 RA Tabata S.;
 RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium Bradyrhizobium japonicum USDA110.";
 RL DNA Res. 9:189-197(2002).
 CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory chain that catalyzes the reduction of oxygen to water. Subunits 1-3 form the functional core of the enzyme complex. CO I is the

CC catalytic subunit of the enzyme. Electrons originating in
 CC cytochrome c are transferred via the copper A center of subunit 2
 CC and heme A of subunit 1 to the binuclear center formed by heme A3
 CC and copper B.
 CC -1- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferricytochrome
 CC c + 2 H(2)O.
 CC -1- PATHWAY: Respiratory chain; terminal step.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- DEVELOPMENTAL STAGE: Free in soil (not as bacteroid).
 CC -1- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X54318; CAA38216.1; -;
 CC EMBL; X54800; CAA38570.1; -;
 CC EMBL; AJ242592; CAB56819.1; -;
 CC EMBL; U33883; AAF78815.1; -;
 CC EMBL; AP005939; BAC46436.1; -;
 CC PIR; S13076; ODZJ1.
 CC HSP; P00396; 20CC.
 CC InterPro: IPR000883; COX1.
 CC Pfam; PF00115; COX1.1.
 CC PRINTS; P01165; CYCOXIDASE1.
 CC PROSITE; PSS0855; COX1.1.
 CC PROSITE; PS00077; COX1_CUB; 1.
 CC Complete proteome; Copper; Heme; Oxidoreductase; Respiratory chain;
 CC Transmembrane.
 CC TRANSMEM 40 60 Potential.
 CC TRANSMEM 87 107 Potential.
 CC TRANSMEM 124 144 Potential.
 CC TRANSMEM 173 193 Potential.
 CC TRANSMEM 211 231 Potential.
 CC TRANSMEM 262 282 Potential.
 CC TRANSMEM 294 314 Potential.
 CC TRANSMEM 332 352 Potential.
 CC TRANSMEM 365 385 Potential.
 CC TRANSMEM 404 424 Potential.
 CC TRANSMEM 441 461 Potential.
 CC TRANSMEM 483 503 Potential.
 CC METAL 85 85 Iron (heme A axial ligand) (Probable).
 CC METAL 268 268 Copper B (probable).
 CC METAL 272 272 Copper B (probable).
 CC METAL 317 317 Copper B (probable).
 CC METAL 318 318 Copper B (probable).
 CC METAL 403 403 Iron (heme A3 axial ligand) (Probable).
 CC METAL 405 405 Iron (heme A axial ligand) (Probable).
 CC METAL 405 405 1'-histidyl-3'-tyrosine (His-Tyr) (By
 CC CROSSLINK 268 272 similarity).
 CC SEQUENCE 541 AA; 59265 MW; 08F9A59D0C90DB85 CRC64;
 CC
 CC Query Match 61.0%; Score 1829; DB 1; Length 541;
 CC Best Local Similarity 61.8%; Pred. No. 3,le-110;
 CC Matches 337; Conservative 72; Mismatches 112; Indels 24; Gaps 4;
 CC
 CC 7 HGHDEHEKQGFETRWMTNKHDKIGLLYLAAGVVGVSIVLFTVYMELEMDPGVQVMCL 66
 CC 15 HGHDEHAHPTGWRVYVSTNKHDKIGTWYLFVAVIAGVIGAMSIAIRAEMLPGVQVIF-- 72
 CC 67 EGARLIADASQCTCTANGHLNVMVTYIGILMNVFVIGIPALPGGNGNLMPLQIGAPDMAF 126
 CC 73 -----HETHYVNFVTSHGLINIFFMVMFAMIGGNGNFMVFMIGAPDMAF 118
 CC 127 PRMNLSFWLFIAGTAMGVASLPAFGGQGLGSGVGVVLYPPLSTR-EAGYSMDLAFV 185
 CC 119 PRMNISFWLPAFGGLLMSFTVEGPGANGVAGTMYVPLSSSGHPGPAVDFAISL 178
 CC 186 HLGASSINGAINMTITFLNVRPQMTHLHKVPLPFSWISIFITAWLILLALPVLGAITWLL 245

```

Db 179 HLAGASSILGAINFITTINMRADGMLTKHNPFLVMSILVTVLELLLSLPLVLAGAITMLL 238
Qy 246 TDRNFGTTFFNPAGGDPILYQHILWFFGHPEVYIIILPGFGIISHVSVFSSFKPKPVFGYL 305
Db 239 TDRNFGTTFFAPDCGGDPVLFQHLWFFGHPEVYIIILPGFGIISHVSVFSSFKPKPVFGYL 298
Qy 306 PMYAMVAIGVLGRVWHAHMYTVGMSLTQOSFPLATWIAVPTGKIFPSWATWNGS 365
Db 299 GMATAMVAIGGIGFVWHAHMYTVGMSLTQOSFPLATWIAVPTGKIFPSWATWNGS 358
Qy 366 VEFKSPMLWAFGFMFLFTVGGVTGIVLAQGLDRVHDYVVAHMYVMSLGAIFAIFA 425
Db 359 IEPEAPMVAIGFIFLFTVGGVTGIVLAQGLDRVHDYVVAHMYVMSLGAIFAIFA 418
Qy 426 GIYFPMKFSGRAPPEAAKLHFWTFPIGANVTFPQHFILGROGMPRRYIDYPEAFALWN 485
Db 419 GWYWFPMKMGYMYNETLAKAHFWTFPIGANVTFPQHFILGROGMPRRYIDYPEAFALWN 478
Qy 486 KVSSYGAFALAFASPLFFIVFVTVLVAAGRETRP---NPMGEFADTLEWTLSPSPPAHTF 542
Db 479 LVSSVGSYISG---FQVILFLICVIDAFKAKVPAGDNPWAGATTLEWTLSPSPPPHQP 534
Qy 543 ETLPK 547
Db 535 EVLPR 539

```

RESULT 15

```

Q36520 PRELIMINARY; PRT; 523 AA.
AC Q36520;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome oxidase subunit 1.
GN Name=coxi;
OS Platymonas subcordiformis.
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Chlorophyta; Prasinophyceae; Platymonas.
CX NCBI_TaxID=3161;
RN [1]
RP MEDLINE=96145517; PubMed=8555450;
RX Kessler U.; Zesche K.;
RT "Physical map and gene organization of the mitochondrial genome from
RT the unicellular green alga Platymonas (Tetraselmis) subcordiformis
RT (Prasinophyceae).";
RL Plant Mol. Biol. 29:1081-1086(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Kessler U.;
RL Thesis (1995), Institute for Plant Physiology, Justus Liebig
RL University.
CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. CO I is the
CC catalytic subunit of the enzyme. Electrons originating in
CC cytochrome c are transferred via the copper A center of subunit 2
CC and heme A of subunit 1 to the binuclear center formed by heme A3
CC and copper B (By similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -1- PATHWAY: Respiratory chain; terminal step.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
DR EMBL; Z47795; CAA87753.1; -
DR PIR; S62706; S62706.
DR HSSP; P00396; 20CC.
DR CO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.

```

```

DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0004129; P:cytochrome-c oxidase activity; IEA.
DR GO; GO:0016491; P:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000983; COX1.
DR Pfam; PF00115; COX1; 1.
DR PRINTS; PRO1165; CYCOXIDASE1.
DR PROSITE; PS00077; COX1; 1.
KW Copper; Electron transport; Heme; Inner membrane; Membrane;
KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
KW Transport.
SQ SEQUENCE 523 AA; 57669 MW; 3E5B17B6E743085C CRC64;

Query Match 59.1%; Score 1773; DB 2; Length 523;
Best Local Similarity 60.1%; Pred. No. 1.3e-106;
Matches 322; Conservative 82; Mismatches 104; Indels 28; Gaps 6;

Qy 17 FFTFMSTNKHDKIGLLYLVAAGVVGFIIVLVYVMELEMDPGVQVMCLGARLIADAS 76
Db 4 PAQRALSTNKHDKIGLLYLVAAGVVGFIIVLVYVMELEMDPGVQVMCLGARLIADAS 76
Qy 77 QTCTANGHLNVMVYTHGILMMFFVGIPALFGSGNYLMLQIGAPDPAFPRNNISFWL 136
Db 49 QVLGNHQLYNYIITAHAFIMIFFMNPALIGGFWNFMICAPDMAPRLNNSFWL 108
Qy 137 FIATMGVASLFAFGGQGLGSGVGVLYPPLST--REAGYSMDLAIFAVHLSGASSIM 194
Db 109 LPPSLLLLLSALV-----EIGAGTGMTVYPPUSSIASHSGASVDLAIFSLHSGASSIL 163
Qy 195 GAINMITTFLNMRAPGMLTKHNPFLVMSILVTVLELLLSLPLVLAGAITMLLTDNFGTTF 254
Db 164 GALNFITILNMRGPGWTHRLPLFVMSVLITAFLLLSLPLVLAGAITMLLTDNFGTTF 223
Qy 255 ENPAGGDPILYQHILWFFGHPEVYIIILPGFGIISHVSVFSSFKPKPVFGYLPMYAVAI 314
Db 224 FDPAGGDPILFQHLWFFGHPEVYIIILPGFGIISHVSVFSSFKPKPVFGYLPMYAVAI 283
Qy 315 GVLGFVWHAHMYTVGMSLTQOSFPLATWIAVPTGKIFSWIATWNGSVFSSFKPKPKMLW 374
Db 284 GILGFIWHAHMYTVGMSLTQOSFPLATWIAVPTGKIFSWIATWNGSVFSSFKPKPKMLF 343
Qy 375 AFGMFLFTVGGVTGIVLAQGLDRVHDYVVAHMYVMSLGAIFAIFAIFAGYFYMPKF 434
Db 344 AIGFLFTVGGVTGIVLAQGLDRVHDYVVAHMYVMSLGAIFAIFAIFAGYFYMPKF 403
Qy 435 SGRAFFPEAAKLHFWTFPIGANVTFPQHFILGROGMPRRYIDYPEAFALWNKYSSYGAFL 494
Db 404 CGLOYSEFLQIHFWLFFVGVNITFFPMHFLGISGMPPRIPDIPDAFAGNAVCVSGYL 463
Qy 495 AFASFLFFIVFVTVLVAAGRETRP---GEFADTLEWTLSPSPPAHTTFLP 546
Db 464 SVLGALFFVYV-VYDLSGEERCCKNPWETVPGTSA-TLEWTLSPSPPAHTTFLP 517

```

Search completed: October 18, 2004, 23:44:31
Job time : 201 secs